מבוא לסטטיסטיקה תאורית ולהסתברות

פרופ' משה חביב, המחלקה לסטטיסטיקה ומדע הנתונים, האוניברסיטה העברית

תודות
רשימות אלו נכתבו ברובן על ידי דנה אוגוסט במהלך קורס שניתן באוניברסיטה העברית. הן נערכו על ידי נחי אברהם.
כמו כן, עזרו לצורך הגרסה הסופית אירנה קפלן, רחל אקסלרוד, שיר משה ואלישבע שורץ כשהאחרונה דאגה בעיקר לקבצי השאלות והפתרונות.
לכולם נתונה תודתי

מארס 2019

תוכן עניינים

VI התפלגות משותפת 115

VII אי־שוויונים 139

חלק I
סטטיסטיקה תיאורית

תפקידה של הסטטיסטיקה התיאורית הוא לעבד נתונים. למשל, נניח שנתונות ההכנסות של כל משקי הבית בישראל. כלומר, נתונים לנו כמה מיליוני מספרים שכל אחד מהם הוא הכנסה של משק בית כלשהו בישראל. אנו נרצה לראות את ה`יער` מתוך ה`עצים`. נרצה לסכם נתונים בתמציתיות או לתאר אותם באמצעות המחשות כמו דיאגרמה או גרף, כך שנקבל תמונה כללית על ההכנסות במשקי בית בישראל.

הסטטיסטיקה התיאורית עוסקת בעיקר (אך לא רק) במשתנים כמותיים. נניח שמתייחסים לגובה של אנשים השייכים לאוכלוסייה מסויימת. נסמן את המשתנה המספרי שמקבל את ערך הגובה של כל אדם ב־Y. נקבע יחידות מדידה קבועות, למשל מטרים, וכל התצפיות (הגבהים של האנשים השייכים לאוכלוסייה) יימדדו באותן יחידות. נניח כי נתונים לנו n אנשים באוכלוסיית היעד, כך שקיימות התצפיות המתאימות $Y_1, Y_2, ..., Y_n$.

הערה: נשים לב כי Y_1 הוא ערכה של התצפית שאנו קוראים לה ראשונה. `ראשונה` איננו יותר מאשר שם או כינוי לצורכי התייחסות. האינדקס אינו מעיד על חשיבותה של התצפית אלא רק על המספר הסידורי שלה בתוך כלל התצפיות. למעשה הוא משמש כשם לצרכי התייחסות. האינדקס מאפשר להתייחס לאוכלוסייה כאל סדרה, כלומר קבוצה שיש בה.

1 מדדי מרכז

נרצה לסכם את התצפיות שקיבלנו במספר או שניים שמייצגים באיזשהו אופן את כלל התצפיות. מספרים אלה מייצגים במידה מסוימת את התכונות של כלל האוכלוסייה הנמדדת, ועל־כן הם מאפשרים להשוות באופן כללי בין אוכלוסיות שונות.

1.1 ממוצע חשבוני (Arithmetic mean או Average)

הגדרה: נניח כי נתונה הסדרה $\{Y_i\}_{i=1}^n$ (כלומר, אוסף התצפיות $Y_1, Y_2, ..., Y_n$).
נאמר שהממוצע החשבוני של הסדרה הוא:

$$\boxed{\overline{Y} = \frac{Y_1 + Y_2 + ... + Y_n}{n} = \frac{1}{n} \cdot \sum_{i=1}^{n} Y_i}$$

דוגמה: נניח שהערכים הנתונים הם $-1, 0, 2, 4, 7$. נשתמש בנוסחה שהגדרנו ונקבל שהממוצע החשבוני הוא:

$$\frac{-1+0+2+4+7}{5} = 2.4$$

מתכונות הממוצע החשבוני

1. **הממוצע החשבוני משמר את יחידות המדידה.**
למשל אם נתונה הסדרה $\{Y_i\}_{i=1}^n$ ביחידות של מטרים, אז גם הממוצע החשבוני $\overline{Y}$ מתקבל ביחידות של מטרים.

2. **הממוצע החשבוני משמר `טרנספורמציה לינארית`.**
עבור כל זוג מספרים קבועים כלשהם a, b מתקיימת הנוסחה:

$$\overline{a \cdot Y + b} = a \cdot \overline{Y} + b$$

נסביר: אם עבור הסדרה $\{Y_i\}_{i=1}^n$ הממוצע החשבוני הוא $\overline{Y}$, אז עבור הסדרה $\{aY_i + b\}_{i=1}^n$ (כלומר, $aY_1+b, aY_2+b, ..., aY_n+b$) הממוצע החשבוני יהיה $a\overline{Y}+b$.
נשים לב שחיבור וכפל בקבוע הן פעולות שנקראות לינאריות.
דוגמה לשימוש היא חישוב הממוצע לאחר שינוי ביחידות המדידה. כך למשל אם נתון ממוצע ביחידות של מטרים, נשתמש בטרנספורמציה הלינארית $f(x) = 100x$ כדי לקבל את הממוצע ביחידות של סנטימטרים.

3. **ממוצע של סכום שווה לסכום הממוצעים.**
עבור כל זוג סדרות $\{Y_i\}_{i=1}^n$, $\{X_i\}_{i=1}^n$ מתקיימת הנוסחה:

$$\overline{X+Y} = \overline{X} + \overline{Y}$$

גם תכונה זו נובעת מהלינאריות של הממוצע החשבוני. שימו לב, כי הסדרות הן בעלות אותו מספר של ערכים.

4. **הממוצע הוא פונקציה שתלויה בכל הערכים.**
אם נשנה ערך אחד מערכי הסדרה – לא משנה איזה ערך – הממוצע בהכרח ישתנה. כמובן, אם שינוי זה הוא כלפי מעלה או מטה, יתקיים שינוי באותו כיוון בממוצע (אך לא באותו ערך).

5. **הממוצע החשבוני מביא למינימום את סכום ריבועי הסטיות של הנתונים ממספר קבוע.**
נסביר: נניח כי נתונה הסדרה $\{Y_i\}_{i=1}^n$ ונניח ש-x הוא מספר כלשהו.
נתבונן בפונקצית ההפסד (או הקנס) שמודדת את סכום ריבועי המרחקים של איברי הסדרה מ-x. נקבל את הפונקציה הבאה:

$$f(x) = (Y_1 - x)^2 + (Y_2 - x)^2 + ... + (Y_n - x)^2 = \sum_{i=1}^{n} (Y_i - x)^2$$

נוכיח שהממוצע החשבוני $\overline{Y}$ הוא המספר x שמביא למינימום את הפונקציה $f(x)$ שהגדרנו. בפרט, לא משנה איזה x נבחר, תמיד יתקיים $f(\overline{Y}) \leq f(x)$.

ראשית נשתמש בנוסחת הכפל הידועה $(a-b)^2 = a^2 - 2ab + b^2$, ונסיק כי מתקיים:

$$f(x) = \sum_{i=1}^{n} (Y_i - x)^2 = (Y_1 - x)^2 + (Y_2 - x)^2 + ... + (Y_n - x)^2$$

$$= \left(Y_1^2 - 2Y_1 x + x^2\right) + \left(Y_2^2 - 2Y_2 x + x^2\right) + ... + \left(Y_n^2 - 2Y_n x + x^2\right)$$

$$= \left(Y_1^2 + Y_2^2 + ... + Y_n^2\right) - 2x\left(Y_1 + Y_2 + ... + Y_n\right) + nx^2$$

$$= \sum_{i=1}^{n} Y_i^2 - 2x \sum_{i=1}^{n} Y_i + nx^2$$

נשים לב שקיבלנו פרבולה `צוחקת` מהצורה $ax^2 + bx + c$, כאשר הקבועים המתאימים במקרה שלנו הם:

$$a = n$$

$$b = -2 \sum_{i=1}^{n} Y_i$$

$$c = \sum_{i=1}^{n} Y_i^2$$

עבור פרבולה צוחקת הנוסחה למציאת הערך $x_{\min}$, כלומר הערך שעבורו הפרבולה $f(x)$ מגיעה למינימום, היא:

$$x_{\min} = \frac{-b}{2a} = \frac{2\sum_{i=1}^{n} Y_i}{2n} = \frac{\sum_{i=1}^{n} Y_i}{n} = \overline{Y}$$

מכאן שהממוצע $\overline{Y}$ הוא הערך שמביא למינימום את פונקציית סכום ריבועי המרחקים של הנתונים מ־x.

6. **הממוצע בריבוע קטן או שווה לממוצע הריבועים.**
במילים אחרות, מתקיים אי השוויון $\overline{Y}^2 \leq \overline{Y^2}$. כמו־כן המקרה של שוויון $\overline{Y}^2 = \overline{Y^2}$ מתקיים אך ורק כאשר כל איברי הסדרה שווים.
הוכחה:

$$0 \leq f(x_{\min}) = \frac{4ac - b^2}{4a} = \frac{4n\sum_{i=1}^{n} Y_i^2 - 4\left(\sum_{i=1}^{n} Y_i\right)^2}{4n} = n\left(\overline{Y^2} - \overline{Y}^2\right)$$

7. **אם נתונים בסדרה k איברים זהים, אז לצורך חישוב הממוצע ניתן להכפיל את האיבר המתאים ב־k.**

כלומר, אם נתונה הסדרה $Y_1, Y_2, \underbrace{Y, ..., Y}_{k\,\text{times}}, ..., Y_n$, אז הממוצע הוא:

$$\overline{Y} = \frac{Y_1 + Y_2 + \overbrace{Y + ... + Y}^{\times k} + ... + Y_n}{n} = \frac{Y_1 + Y_2 + kY + ... + Y_n}{n}$$

1.2 ממוצע הנדסי (Geometric mean)

הגדרה: נניח כי נתונה הסדרה $\{Y_i\}_{i=1}^n$ שכל איבריה אי־שליליים.
נאמר שהממוצע ההנדסי (או הגאומטרי) של הסדרה הוא:

$$\boxed{x = \sqrt[n]{Y_1 \cdot Y_2 \cdot ... \cdot Y_n}}$$

ראינו כי הממוצע החשבוני ממזער פונקצית קנס מסוימת. נראה כי הממוצע הגאומטרי ממזער פונקצית קנס אחרת. נתבונן בסדרה שבה מוציאים לוגריתם מכל אחד מהאיברים.
נקבל את הסדרה $\{\log Y_i\}_{i=1}^n$.
נגדיר את פונקצית ההפסד $g(x)$ באופן הבא:

$$g(x) = (\log Y_1 - \log x)^2 + (\log Y_2 - \log x)^2 + ...$$

$$+ (\log Y_n - \log x)^2 = \sum_{i=1}^n (\log Y_i - \log x)^2$$

נשים לב שמהדיון לעיל בו ראינו שממוצע הוא הערך שממזער את פונקציית סכום ריבועי המרחקים, נובע שהערך שממזער את הפונקציה $g(x)$ הוא הערך של x עבורו $\log x = \overline{\log Y}$.
לפי חוקי הלוגריתמים:

$$\overline{\log Y} = \frac{1}{n}(\log Y_1 + \log Y_2 + ... + \log Y_n) = \frac{1}{n}\log(Y_1 \cdot Y_2 \cdot ... \cdot Y_n) =$$

$$= \log\left((Y_1 \cdot Y_2 \cdot ... \cdot Y_n)^{\frac{1}{n}}\right) = \log \sqrt[n]{Y_1 \cdot Y_2 \cdot ... \cdot Y_n}$$

נסיק מכך:[1]

$$\log x = \overline{\log Y} = \log \sqrt[n]{Y_1 \cdot Y_2 \cdot ... \cdot Y_n}$$

$$\Downarrow$$

$$x = \sqrt[n]{Y_1 \cdot Y_2 \cdot ... \cdot Y_n}$$

מכאן שבדומה לממוצע החשבוני שהוגדר כמספר שממזער את סכום ריבועי המרחקים של איברי הסדרה ממנו, הממוצע ההנדסי מוגדר כמספר שהלוגריתם שלו ממזער את סכום ריבועי המרחקים של לוגריתם איברי הסדרה ממנו.

[1] פונקציית הלוגריתם היא פונקציה הפיכה, ולכן ניתן לצמצם אותה משני הצדדים.

מתכונות הממוצע ההנדסי

1. הממוצע ההנדסי משמר את יחידות המדידה.
2. לכל $a \geq 0$ מתקיים:

$$\sqrt[n]{aY_1 \cdot aY_2 \cdot ... \cdot aY_n} = a\sqrt[n]{Y_1 \cdot Y_2 \cdot ... \cdot Y_n}$$

(כלומר, הממוצע ההנדסי לינארי ביחס לכפל בקבוע).

3. לכל $b \neq 0$ מתקיים:

$$\sqrt[n]{(Y_1+b)\cdot(Y_2+b)\cdot ... \cdot(Y_n+b)} \neq \sqrt[n]{Y_1 \cdot Y_2 \cdot ... \cdot Y_n} + b$$

(כלומר, הממוצע ההנדסי אינו לינארי ביחס לחיבור בקבוע).

4. לכל סדרה מהצורה $\{x_j\}_{j=1}^n$ מתקיים:

$$\sqrt[n]{x_1Y_1 \cdot x_2Y_2 \cdot ... \cdot x_nY_n} = \sqrt[n]{x_1 \cdot x_2 \cdot ... \cdot x_n} \cdot \sqrt[n]{Y_1 \cdot Y_2 \cdot ... \cdot Y_n}$$

1.3 ממוצע הרמוני (Harmonic mean)

הגדרה: נניח כי נתונות הסדרה $\{Y_i\}_{i=1}^n$, שכל איבריה שונים מ-0.
נאמר שהממוצע ההרמוני של הסדרה הוא:

$$\boxed{x = \frac{n}{\frac{1}{Y_1}+\frac{1}{Y_2}+...+\frac{1}{Y_n}} = \frac{1}{\overline{\frac{1}{Y}}}}$$

במילים: זהו המספר ההופכי לממוצע החשבוני של הופכי אברי הסדרה

נשים לב שבתרגום למונחי פונקציית הפסד, נחפש x שימזער את הפונקציה:

$$\sum_{i=1}^{n}\left(\frac{1}{Y_i}-\frac{1}{x}\right)^2$$

אם נבחר $\frac{1}{x} = \overline{\frac{1}{Y}}$ נקבל מינימום של הפונקציה, כפי שהסברנו לעיל בנוגע לממוצע החשבוני.
ולכן נציב בשוויון הנדרש את התוצאה ונקבל:

$$x = \frac{1}{\overline{\frac{1}{Y}}} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{Y_i}} = \frac{n}{\sum_{i=1}^{n}\frac{1}{Y_i}}$$

1.4 בחירת סוג הממוצע

אי־שוויון הממוצעים: ראשית נזכיר (מבלי להוכיח) תוצאה ידועה של המתמטיקאי אוגוסטין קושי, שנקראת `אי־שוויון הממוצעים`.

משפט זה קובע שלכל סדרה של מספרים חיוביים $\{x_i\}_{i=1}^{n}$, מתקיים עבור שלושת הממוצעים שהגדרנו אי השוויון הבא:

$$\boxed{\frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+...+\frac{1}{x_n}} \leq \sqrt[n]{x_1 \cdot x_2 \cdot ... \cdot x_n} \leq \frac{x_1+x_2+...+x_n}{n}}$$

המשפט קובע ששוויון מתקיים אם ורק אם כל המספרים בסדרה זהים.

דוגמה 1

ברוקר נוכח שבשלוש שנים עוקבות מניה הכפילה את עצמה בערכים $1.1,\ 1.23,\ 0.9$. לכן שווי המניה לאחר שלוש השנים הוכפל בערך של $1.1 \cdot 1.23 \cdot 0.9 = 1.2177$. בממוצע חשבוני, ערך המניה הוכפל בשנה בערך של $\frac{1.1+1.23+0.9}{3} = 1.077$. נשים לב שאם כל שנה היינו מכפילים את ערך המניה ב–1.077 היינו מרוויחים יותר ממה שהושג במציאות: $1.077^3 = 1.248$.

כעת נמצא תשואה קבועה כזאת שב–3 שנים תניב את הרווח שהתקבל בפועל (1.2177):

$$x^3 = 1.2177 \Rightarrow x = \sqrt[3]{1.2177} = 1.068$$

נשים לב שזה התקבל למעשה באמצעות חישוב הממוצע ההנדסי. מכאן שבחישוב תשואות הממוצע ההנדסי הוא מדד מרכזי מהימן יותר מאשר הממוצע החשבוני.

הערה: בהתאם לאי־שוויון הממוצעים, נוכחנו שהממוצע ההנדסי קטן מהממוצע החשבוני.

דוגמה 2

מכונית נוסעת מחיפה לתל־אביב, מרחק של 100 ק`מ, במהירות קבועה של 100 קמ`ש. דרך זו כמובן תארך שעה. המכונית חוזרת את אותה הדרך במהירות של 50 קמ`ש, וכעת הדרך תארך כבר שעתיים. מהי המהירות הממוצעת של המכונית?

תשובה פזיזה עלולה לקבוע שהמהירות הממוצעת היא 75 קמ`ש. אולם לכאורה זו תשובה שגויה, כי מהירות ממוצעת מוגדרת כסך המרחק חלקי סך הזמן ולכן הממוצע הוא $\frac{200}{3} = 66.667$. ואכן אם היינו דוגמים את מהירות המכונית בכל דקה היינו מקבלים את הנתונים, $\{100\}_{j=1}^{60}$, $\{50\}_{i=1}^{120}$, כך שהממוצע החשבוני בכל הדקות הוא:

$$\frac{60 \cdot 100 + 120 \cdot 50}{180} = 66.667$$

לעומת זאת נשים לב שאם היינו בודקים את מהירות המכונית בכל קילומטר היינו מקבלים $\{100\}_{j=1}^{100}$, $\{50\}_{i=1}^{100}$, כך שהממוצע החשבוני הוא:

$$\frac{100 \cdot 100 + 100 \cdot 50}{200} = 75$$

אם כן מהי התשובה הנכונה? אין תשובה נכונה יחידה. ממוצע חייב להתייחס ליחידות המדידה שבהן אנו בוחרים למדוד. במקרה זה עלינו להחליט האם מעוניינים לבדוק מהירות ממוצעת לדקה (זמן) או מהירות ממוצעת לקילומטר (מרחק). כמובן שהבחירה הראשונה היא זו המעשית יותר.

הערה: עד כה ניגשנו להגדיר מדדי מרכז או ממוצעים למיניהם כדי שימזערו פונקציות הפסד מסוימות. קיימת גישה אחרת להגדיר את הממוצעים, לפיה מעוניינים להחליף את כל איברי הסדרה במספר קבוע שיביא אותנו לאותו מקום.

- **ממוצע חשבוני:** אם נבדוק מהו הקבוע c המקיים $\sum_{i=1}^{n} X_i = \underbrace{c + c + ... + c}_{n\,\text{times}}$, נקבל את הממוצע החשבוני $c = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.
- **ממוצע הנדסי:** אם נבדוק מהו הקבוע c המקיים $\prod_{i=1}^{n} X_i = \underbrace{c \cdot c \cdot ... \cdot c}_{n\,\text{times}}$, נקבל את הממוצע ההנדסי $c = \sqrt[n]{\prod_{i=1}^{n} X_i}$ (עבור איברים שכולם חיוביים).
- **ממוצע הרמוני:** אם נבדוק מהו הקבוע c המקיים $\sum_{i=1}^{n} \frac{1}{X_i} = \underbrace{\frac{1}{c} + \frac{1}{c} + ... + \frac{1}{c}}_{n\,\text{times}}$, נקבל את הממוצע ההרמוני $c = \frac{n}{\sum_{i=1}^{n} X_i}$.

דוגמה 3

נניח כי עוברים מרחק בגודל a מספר כלשהו של פעמים שנסמן n. בכל אחת מהפעמים עוברים את המרחק במהירות Y_i. כלומר נתונה לנו סדרת המהירויות בכל פעם: $\{Y_i\}_{i=1}^{n}$. נשים לב שמשך הזמן שאורכת הדרך בפעם ה-i הוא $\frac{a}{Y_i}$, ולכן הזמן שאורך לעבור את המרחק הכולל של na הוא $a\sum_{i=1}^{n}\frac{1}{Y_i}$.

שאלה: מהי המהירות הקבועה (שנסמן כנעלם x) שאם נשתמש בה בכל n הפעמים, סך הזמן שיארך לעבור את המרחק na יהיה שווה לסך הזמן שאורך לעבור את אותו המרחק במהירויות $\{Y_i\}_{i=1}^{n}$ בהתאמה?

אם נתרגם את השאלה לסימונים בהם השתמשנו, נחפש x שיקיים את השוויון:

$$a\sum_{i=1}^{n}\frac{1}{Y_i} = \frac{na}{x}$$
$$\Updownarrow$$
$$x = \frac{n}{\sum_{i=1}^{n}\frac{1}{Y_i}}$$

וזו בדיוק הגדרת הממוצע ההרמוני.

1.5 חציון (median)

הגדרה: נניח כי נתונה הסדרה $\{Y_i\}_{i=1}^{n}$ נאמר שהחציון של הסדרה הוא הערך האמצעי של הערכים, והוא מסומן ב-$med\,(Y)$. בפרט, אם מספר התצפיות הוא אי־זוגי, מסדרים את הערכים שבסדרה בסדר עולה (או יורד) ובוחרים את הערך האמצעי. אם מספר התצפיות הוא אי־זוגי, החציון מוגדר להיות כל ערך שנמצא בין שני הערכים האמצעיים של הסדרה[2].

דוגמה: נניח שנתונות התצפיות $Y = \{-1, 0, 2, 4, 7\}$, אז $med\,(Y) = 2$.

נראה שגם החציון ממזער פונקציית הפסד כלשהי.
נגדיר פונקציית את פונקצית ההפסד $h\,(x)$ באופן הבא:

$$h\,(x) = |Y_1 - x| + |Y_2 - x| + ... + |Y_n - x| = \sum_{i=1}^{n} |Y_i - x|$$

נשים לב שמתקיים:

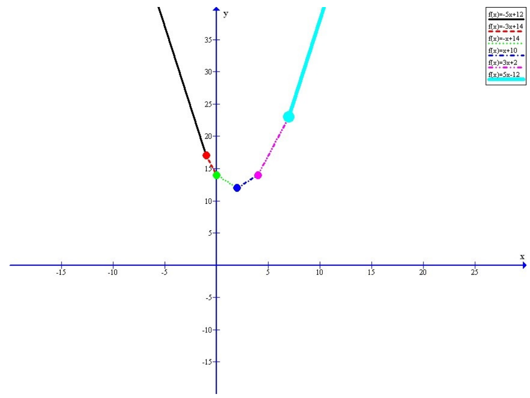
$$h\,(x) = \sum_{i=1}^{n} |Y_i - x| = n \cdot \frac{1}{n} \cdot \sum_{i=1}^{n} |Y_i - x| = n \cdot \overline{|Y - x|}$$

בדוגמה שהזכרנו, הפונקציה המתקבלת היא:

$$h\,(x) = |-1-x| + |x| + |2-x| + |4-x| + |7-x| = \begin{cases} -5x + 12 & \text{for } x \le -1 \\ -3x + 14 & \text{for } -1 \le x \le 0 \\ -x + 10 & \text{for } 0 \le x \le 2 \\ x + 10 & \text{for } 2 \le x \le 4 \\ 3x + 2 & \text{for } 4 \le x \le 7 \\ 5x - 12 & \text{for } 7 \le x \end{cases}$$

נסביר: נשים לב שעבור $x \le -1$ כל הביטויים בתוך סימני הערך המוחלט הם חיוביים, ולכן בתחום זה ניתן לוותר על סימנים אלו, ואז הפונקציה מוגדרת $h\,(x) = -5x + 12$. כאשר $-1 \le x \le 0$ עובדה זו נכונה רק עבור ארבעת המחוברים האחרונים. עבור המחובר הראשון הערך המוחלט הופך את סימנו של הביטוי הרשום בתוכו והוא $x + 1$, לכן עבור תחום זה הפונקציה מוגדרת $h\,(x) = -3x + 14$. בתחום $0 \le x \le 2$ שני המחוברים הראשונים הופכים סימן ושלושת האחרונים לא, ולכן עבור תחום זה הפונקציה מוגדרת $h\,(x) = -x + 10$. וכן הלאה.
הגרף של פונקציה זו הוא:

[2] לעיתים מקובל לקחת את הממוצע החשבוני בין שני ערכים אמצעיים אלה



ניתן לראות כי עבור $x < 2$ הפונקציה מונוטונית יורדת, ועבור $2 \leq x$ היא מונוטונית עולה. בפרט, הנקודה $x = 2$ היא נקודת מינימום שבה הפונקציה משנה את כיוון המונוטוניות. נשים לב ש-$med\left(Y\right) = 2$. כלומר, פונקציית ההפסד שהגדרנו מתמזערת בנקודת החציון. במקרה שבו קיים מספר זוגי של ערכים בסדרה, כל הקטע שבין שני הערכים האמצעיים יהיה בשיפוע 0 וכל הנקודות שבו יהוו מינימום של פונקציית ההפסד שהגדרנו.

נתבונן שוב בפונקציית ההפסד שהגדרנו $h\left(med(y)\right) = \sum_{i-=1}^{n} |Y_i - \overline{Y}|$. ניתן להבחין שעבור מחצית הערכים הגדולים מהחציון הסימן אינו משתנה כי מתקבל מספר חיובי, ועבור מחצית הערכים הקטנים מהחציון הסימן משתנה משלילי לחיובי. לכן נוכל להסיק שערך פונקציית ההפסד בחציון הוא ההפרש בין סכום המחצית העליונה של הערכים לבין סכום המחצית התחתונה שלהם. כלומר, אם נתונה הסדרה $Y_1 \leq Y_2 \leq \ldots \leq Y_n$, אז מתקיים:

$$h\left(med\left(Y\right)\right) = \begin{cases} \sum\limits_{i=\frac{n+1}{2}}^{n} Y_i - \sum\limits_{i=1}^{\frac{n-1}{2}} Y_i & \text{if } n \text{ is odd} \\ \sum\limits_{i=\frac{n}{2}+1}^{n} Y_i - \sum\limits_{i=1}^{\frac{n}{2}} Y_i & \text{if } n \text{ is even} \end{cases}$$

מתכונות החציון

1. החציון משמר את יחידות המדידה.
2. לכל a, b קבועים מתקיימת הנוסחה:

$$med\left(aY + b\right) = a \cdot med\left(Y\right) + b$$

3.

$$med\left(X + Y\right) \neq med\left(X\right) + med\left(Y\right)$$

נשים לב שחיבור קבוצות מהצורה שסימנו "$X+Y$" הוא חיבור של זוגות איברים בעלי אותו אינדקס. נתבונן למשל בסדרות הבאות:

$$med\,(1,2,4)=2$$

$$med\,(1,3,2)=2$$

$$med\,(1+1,2+3,4+2)=med\,(2,5,6)=5$$

4. החציון אינו רגיש כמו הממוצע לשינוי בערכי הסדרה. שינוי של ערך בסדרה יכול להשפיע על החציון רק אם הערך היה גדול (קטן) מהחציון ושונה להיות קטן (גדול) ממנו.

הרחבה: נדון בדוגמה מעניינת שתשקף יתרון לשימוש בחציון כמדד מרכזי על־פני הממוצע.

מספר בתים ממוקמים לאורך רחוב אחד, נניח במקומות $x_1 \leq x_2 \leq ... \leq x_n$. מתכנן מרכזי רוצה למקם תחנת אוטובוס אחת ברחוב. סביר שימקם אותה היכן שהוא 'באמצע', זאת אומרת למשל בממוצע החשבוני בין ה־x־ים או בחציון. נניח כעת כי המתכנן אינו יודע את מיקום הבתים והוא נאלץ להעזר באינפורמציה שיתנו לו הדיירים. הדיירים מצידם אינם חייבים לומר את האמת כאשר ברצונם כי התחנה תמוקם קרוב ככל הניתן לביתם. נראה ראשית כי אם התחנה ממוקמת בממוצע החשבוני לפי הדיווחים, אין זה בכלל ברור כי כולם יאמרו את האמת. למשל, `נביא` מביניהם שיודע מה האחרים ידווחו תמיד יוכל לדווח דבר שקר ולמקם את התחנה במרחק אפס מביתו[א]. נכון, אפסו נביאים, אבל מטרת דוגמא זו היא להראות כי אמירת אמת אינה מובנת מאליה. לעומת זאת, אם המיקום יהיה בחציון המקומות המדווחים אין לאף אחד עניין לדווח שקר. במילים אחרות, לא משנה מה יעשו האחרים לומר את האמת זו האסטרטגיה הטובה ביותר (אם כי לא בהכרח באופן יחיד). ניתן לראות זאת כך. נניח כי לאחר דיווח אמת x_i על ידי דייר i ודיווח כלשהו של האחרים, נקבע החציון להיות קטן מ־ x_i. האם כדאי לדייר זה לשנות את דיווחו? אם יגדיל את דיווחו החציון לא ישתנה, ואם יקטין אותו, החציון או שלא ישתנה או שיוקטן עוד יותר וכך רק ירחיק עוד יותר את החציון מ־x_i. שיקול דומה קיים גם למקרה שהחציון לפי הדיווחים יצא גדול מ־x_i.

[א] אם כולם פרט ל־i ידווחו y_j, $1 \leq j \neq i \leq n$ הדייר ה־i ידווח על y_i כך שיתקבל $x_i = \frac{\sum_{i=1}^{n} y_i}{n}$ זאת אומרת, $y_i = n \cdot x_i - \sum_{j \neq i}^{n} y_j$

2 מדדי פיזור

המידע שטמון במדדים מרכזיים מתעלם מהפיזור של הערכים סביב אותו מדד מרכזי. כך למשל הממוצע של $\{4,5,6\}$ והממוצע של $\{0,5,10\}$ שניהם שווים ל־5, על אף שהערכים בסדרה הראשונה מפוזרים במרחקים גדולים יותר מהממוצע.
נחפש מדדים שנכנה מדדי פיזור, שייתנו לנו מידע אודות מידת הפיזור של הערכים סביב המדד המרכזי. פיזור של אוכלוסייה כלשהי תמיד ייקבע ביחס למדד מרכזי כלשהו של האוכלוסייה הרלוונטית. נציע מדד שנראה טבעי (אך למעשה שגוי) למדידת הפיזור של ערכים

סביב הממוצע: נניח כי נתונה הסדרה $\{Y_i\}_{i=1}^n$ כך ש-$\overline{Y}$ הוא הממוצע שלה. נגדיר מדד פיזור כממוצע הסטיות בין האיברים לבין הממוצע. כלומר הממוצע של הסדרה $\left\{Y_i-\overline{Y}\right\}_{i=1}^n$:

$$\overline{Y-\overline{Y}}=\frac{\left(Y_1-\overline{Y}\right)+\left(Y_2-\overline{Y}\right)+...+\left(Y_n-\overline{Y}\right)}{n}=\frac{1}{n}\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)$$

נזכור שהראינו כי $\overline{aY+b}=a\overline{Y}+b$, ולכן נסיק:[3]

$$\overline{Y-\overline{Y}}=\overline{Y}-\overline{Y}=0$$

לכן נסיק שמדובר במדד פיזור חסר משמעות, כי הוא קבוע ושווה ל-0.

2.1 שונות (Variance)

הגדרה: נניח כי נתונה הסדרה $\{Y_i\}_{i=1}^n$ כך ש-$\overline{Y}$ הוא הממוצע שלה. נגדיר את השונות של הסדרה כממוצע של ריבועי הסטיות מהממוצע - $\overline{\left(Y_i-\overline{Y}\right)^2}$. כלומר הממוצע של הסדרה $\left\{\left(Y_i-\overline{Y}\right)^2\right\}_{i=1}^n$:

$$\boxed{Var\left(Y\right)=\frac{\left(Y_1-\overline{Y}\right)^2+\left(Y_2-\overline{Y}\right)^2+...+\left(Y_n-\overline{Y}\right)^2}{n}=\frac{1}{n}\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2}$$

נוסחה: נשים לב שמתקיים:

$$\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2=\sum_{i=1}^{n}\left(Y_i^2-2Y_i\overline{Y}+\overline{Y}^2\right)=\sum_{i=1}^{n}Y_i^2-2\overline{Y}\sum_{i=1}^{n}Y_i+n\overline{Y}^2=$$

$$=n\cdot\frac{1}{n}\left(\sum_{i=1}^{n}Y_i^2-2\overline{Y}\sum_{i=1}^{n}Y_i+n\overline{Y}^2\right)=n\overline{Y^2}-2n\overline{Y}^2+n\overline{Y}^2=n\left(\overline{Y^2}-\overline{Y}^2\right)$$

ומכאן נובעת נוסחה עבור השונות:

$$\boxed{Var\left(Y\right)=\overline{Y^2}-\overline{Y}^2}$$

מנוסחה זו ניתן להסיק שוב שמתקיים $\overline{Y^2}\geq\overline{Y}^2$, שכן תמיד $Var\left(Y\right)\geq 0$. (ראה עמוד 8)

מתכונות השונות

1. השונות מתקבלת ביחידות מדידה שהן ריבוע של יחידות המדידה של ערכי הסדרה. למשל, אם ערכי הסדרה נמדדים ביחידות של מטר, השונות מתקבלת ביחידות של מטר רבוע.

---

[3] נשים לב ש-$\overline{\overline{Y}}=\overline{Y}$

2. השונות אדישה לחיבור בקבוע:

$$\mathrm{Var}\left(Y+b\right)=\mathrm{Var}\left(Y\right)$$

ההוכחה לכך פשוטה:

$$\mathrm{Var}\left(Y+b\right)=\frac{1}{n}\sum_{i=1}^{n}\left(Y_i+b-\left(\overline{Y+b}\right)\right)^2$$

$$=\frac{1}{n}\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2=\mathrm{Var}\left(Y\right)$$

3. עבור הכפלה בקבוע מתקיים:

$$\mathrm{Var}\left(aY\right)=a^2\mathrm{Var}\left(Y\right)$$

הוכחה:

$$\mathrm{Var}\left(aY\right)=\frac{1}{n}\sum_{i=1}^{n}\left(aY_i-\overline{aY}\right)^2=\frac{1}{n}\sum_{i=1}^{n}\left(aY_i-a\overline{Y}\right)^2=$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(a\left(Y_i-\overline{Y}\right)\right)^2=a^2\frac{1}{n}\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2=a^2\mathrm{Var}\left(Y\right)$$

נקבל משתי התוצאות הקודמות שלכל a,b קבועים מתקיים:

$$\mathrm{Var}\left(aY+b\right)=a^2\mathrm{Var}\left(Y\right)$$

■

2.2 סטיית תקן (Standard deviation)

הגדרה: נניח כי נתונה הסדרה $\{Y_i\}_{i=1}^{n}$. סטיית התקן של Y היא:

$$\mathrm{SD}\left(Y\right)=\sqrt{\mathrm{Var}\left(Y\right)}$$

תכונות סטיית התקן

1. סטיית התקן משמרת יחידות מדידה. למשל, אם ערכי הסדרה נמדדים במטרים, גם סטיית התקן תמדד במטרים.
2. כמו השונות, סטיית התקן אדישה לחיבור קבוע:

$$\mathrm{SD}\left(Y+b\right)=\mathrm{SD}\left(Y\right)$$

3. עבור הכפלה בקבוע מתקיים:

$$\mathrm{SD}\,(aY) = |a|\,\mathrm{SD}\,(Y)$$

נשים לב: $\sqrt{a^2} = |a|$.
נקבל משתי התוצאות הקודמות שלכל a, b קבועים מתקיים:

$$\mathrm{SD}\,(aY + b) = |a|\,\mathrm{SD}\,(Y)$$

2.2.1 אי־שוויון צ'בישב

אי־שוויון צ'בישב קובע שבכל סדרת תצפיות Y עבור כל $k > 0$, לפחות $1 - \frac{1}{k^2}$ מהתצפיות נופל במרחק של עד $\pm k \cdot SD\,(Y)$ מהממוצע. למשל עבור $k = 3$, לפחות $1 - \frac{1}{9} = \frac{8}{9}$ מהתצפיות נמצאות במרחק של עד 3 סטיות־תקן (למעלה או למטה) מהממוצע. אי־שוויון זה מעניק משמעות פורמלית לטענה שלא ייתכן שחלק גדול מידי מהאוכלוסייה נמצא במרחק רב מידי מהממוצע, כאשר היחידות בהן נמדד המרחק הן מספר סטיות התקן של הערך מממוצע הסדרה $\overline{Y}$. נשים לב שמרחק זה יכול להיות שלילי. הוכחה תינתן בהמשך (עמוד 140)

דוגמא:

מקרה פרטי מעניין הוא זה שבו הערכים היחידים שבאים בחשבון הם 0 או 1. הם באים לציין קיומה או אי קיומה של תכונה בפריט מסויים באוכלוסייה. בפרט, הספרה 1 מציינת את קיום התכונה באותו פריט ו־0 אחרת. במקרה זה $\bar{X}$ מציין את השכיחות היחסית של התכונה באוכלוסיה.
ישנו קשר מידי בין $\bar{X}$ ובין $Var(X)$ במקרה זה:
<u>טענה:</u> $Var(X) = \bar{X}(1-\bar{X})$ ואז $SD(X) = \sqrt{\bar{X}(1-\bar{X})}$
<u>הוכחה:</u> נשים לב כי למקרה שבו $x_i = 1$ או $x_i = 0$, מתקיים $x_i^2 = x_i$. ולכן,

$$Var(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X})^2 = \frac{1}{n}\left(\sum_{i=1}^{n} x_i^2 - 2\bar{X}\sum_{i=1}^{n} x_i + n\bar{X}^2\right) =$$

$$= \frac{1}{n}\left(\sum_{i=1}^{n} x_i - 2n\bar{X}^2 + n\bar{X}^2\right) = \frac{1}{n}(n\bar{X} - 2n\bar{X}^2 + n\bar{X}^2) = \bar{X} - \bar{X}^2 = \bar{X}(1-\bar{X})$$

כנדרש.

2.3 ציוני־תקן (תיקנון)

הגדרה: נניח שנתונה הסדרה $\{Y_i\}_{i=1}^{n}$, כך ש־$\overline{Y}$ הוא הממוצע שלה ו־$\mathrm{SD}\,(Y)$ היא סטיית התקן. נגדיר את סדרת ציוני התקן באופן הבא:

$$\{Z_i\}_{i=1}^{n}=\left\{\frac{Y_i-\overline{Y}}{\mathrm{SD}\left(Y\right)}\right\}_{i=1}^{n}$$

טרנספורמציה לינארית זו שביצענו (כפלנו ב-$\frac{1}{\mathrm{SD}(Y)}$ וחיברנו $\frac{-\overline{Y}}{\mathrm{SD}(Y)}$) נקראת `תיקנון` של הסדרה.

ציוני התקן אינם תלויים ביחידות המדידה המקוריות. כך למשל נתבונן בציון־תקן של סדרה ביחידות של ס`מ וביחידות של מטרים:

$$\frac{100Y_i-\overline{100Y}}{\mathrm{SD}\left(100Y\right)}=\frac{100Y_i-100\overline{Y}}{|100|\,\mathrm{SD}\left(Y\right)}=\frac{Y_i-\overline{Y}}{\mathrm{SD}\left(Y\right)}$$

הסבר: נניח שנתונה אוכלוסייה של אנשים ולכל אחד מהם נתון ה-IQ שלו. קיבלנו מידע שה-IQ של אדם מסוים גבוה מהממוצע. מידע זה מקבל משנה חשיבות אם ידוע שהפיזור סביב הממוצע הוא קטן, יותר מאשר במצב שבו הפיזור רב. במצב שבו הפיזור קטן סביב הממוצע `קשה` יותר להתרחק מהממוצע, ולכן IQ גבוה במקרה זה מהווה תופעה משמעותית יותר מאשר במקרה האחר. כדי להעניק חשיבות לעובדה שאדם זה מעל לממוצע תוך התחשבות במידת הפיזור, נתקנן את ה-IQ שלו.

היבט נוסף בו התיקנון שימושי, הוא מצב בו מעוניינים להשוות בין פרטים שונים באוכלוסיות שונות. למשל השוואה בין גובהם היחסי של שחקן ושחקנית כדורסל. כמו כן תיקנון שימושי במצב בו מעוניינים להשוות בין פרטים הנמדדים ביחידות מידה שונות. למשל האם אדם מסוים הוא גבוה יותר או כבד יותר.

תכונות ציוני התקן

1. הממוצע של ציוני תקן הוא 0.
הוכחה:

$$\overline{Z}=\overline{\frac{Y}{\mathrm{SD}\left(Y\right)}-\frac{\overline{Y}}{\mathrm{SD}\left(Y\right)}}=\frac{\overline{Y}}{\mathrm{SD}\left(Y\right)}-\frac{\overline{Y}}{\mathrm{SD}\left(Y\right)}=0$$

2. השונות של ציוני תקן היא 1, ולפיכך גם סטיית התקן היא 1.
הוכחה:

$$\mathrm{Var}\left(Z\right)=\mathrm{Var}\left(\tfrac{Y}{\mathrm{SD}(Y)}-\tfrac{\overline{Y}}{\mathrm{SD}(Y)}\right)=\mathrm{Var}\left(\tfrac{Y}{\mathrm{SD}(Y)}\right)$$

$$=\tfrac{1}{\mathrm{SD}(Y)^2}\mathrm{Var}\left(Y\right)=\tfrac{\mathrm{Var}(Y)}{\mathrm{Var}(Y)}=1$$

2.4 היסטוגרמה

היסטוגרמה היא שיטה להצגה של נתונים מרובים.

כדי ליצור היסטוגרמה עבור אוסף נתון של תצפיות נעשה את הדברים הבאים:

1. נקבע **טווחים** של ערכים שבכל אחד מהם ייפלו כמה תצפיות (הטווחים יכולים להיות שונים זה מזה באורכם), **ונמספר** כל אחת מקבוצות התצפיות שבטווחים. קביעת הטווחים היא משימה מורכבת לעתים ולא חד־משמעית, כי מצד אחד חלוקה לטווחים מצומצמים (כלומר לקבוצות רבות של ערכים) מעניקה מידע יותר מדויק, אבל מאידך מטשטשת את התמונה הכללית כי מספר התצפיות בכל קטע יהיה קטן.

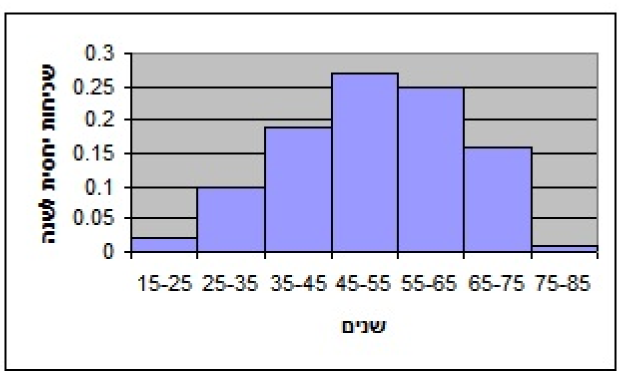
2. בשלב הבא נחשב את **השכיחות היחסית** של כל אחת מקבוצות הערכים. כלומר נבדוק מהו השיעור של כל קבוצה מתוך כלל התצפיות.

3. **נשרטט גרף** לפי השכיחויות היחסיות בשיטה הבאה:
– ניצור מערכת צירים שעל ציר ה־x יחידות המדידה של ערכי התצפיות. ביחידות המדידה של ציר ה־y נדון בהמשך.
– נקבע שרירותית יחידת שטח כללית על המישור, וגודלה של יחידה זו יוגדר 1 וייצג את כלל האוכלוסיה.
– נשרטט מלבן לכל קבוצה. רוחב המלבן (על ציר ה־x) ייקבע לפי הטווח המתאים לקבוצה, וגובה המלבן שמכונה **צפיפות** (על ציר ה־y) יהיה השכיחות היחסית של הקבוצה חלקי אורכה. ובפרט, שטח המלבן יהיה השכיחות היחסית של הקבוצה.

דוגמה: נתונים 100 אנשים שהתפלגות הגילאים שלהם היא:

שכיחות יחסית	שכיחות	גיל
0.02	2	15–25
0.1	10	25–35
0.19	19	35–45
0.27	27	45–55
0.25	25	55–65
0.16	16	65–75
0.01	1	75–85
1	**100**	סך הכל

היסטוגרמה של נתונים אלה תיראה כך:



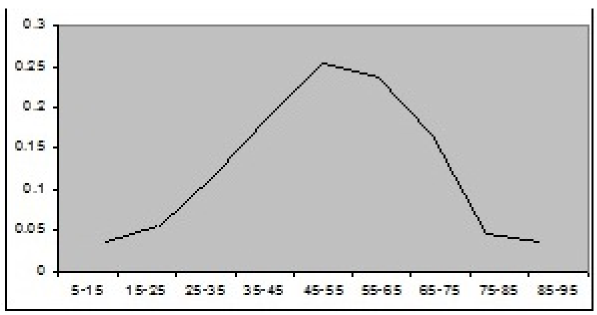
ניתן לראות לפי הדוגמא זו כי כאשר אורכי הקטעים שונים, גבהי המלבנים גם הם פרופורציונלים לשכיחות היחסית של הקבוצה.

הסבר על מושג הצפיפות

עד כה עסקנו ביחידות ציר ה–x ובשכיחות היחסית שמייצג שטח המלבן. היחידות שעל ציר ה–y יוגדרו כצפיפות. מקור המונח הוא שגובה המלבן מגדיר את קצב צבירת השכיחות היחסית, ליחידת x. הצפיפות לא מייצגת שטח כמובן, והיא גם לא השכיחות היחסית. הצפיפות היא השכיחות היחסית חלקי יחידות המדידה של x. כך למשל גובה מלבן של 0.19 אומר שכל פעם שנתקדם יחידה לאורך ציר ה–x, נצבור עוד 1.9% מהשכיחות היחסית.
נציין כי בספרים ותוכנות שונים, לעיתים יחידות המדידה של ציר ה–y מוגדרות כ`הסתברות` או `שכיחות יחסית`. זה רחוק מלהיות נכון. למשל, צפיפות יכולה להיות גדולה בערכה מ–1 בערכים נומרים. כאשר רוחבי הקטעים שווים, הערכים הרשומים על ציר ה–y הם פרופורציונליים לשכיחות היחסית של הקטעים הרלוונטיים, אך אין הם ביחידות המדידה שלו, כנדרש מכל פונקציה.

פוליגון

פוליגון הוא מצולע שקודקודיו הם אמצעי הפאה העליונה של המלבנים בהיסטוגרמה. בדוגמה הקודמת הפוליגון המתאים הוא:



3 מדדי קשר בין משתנים

לאחר שעסקנו בכל משתנה בנפרד נרצה לאפיין קשר בין משתנים שונים וכיצד הם נעים יחד. למשל, האם ניתן ללמוד משהו על ערכו של האחד אם ידוע ערכו של השני? האם ניתן לקבוע בהכללה שכאשר ערכו של אחד גדל כך גם השני? או להיפך? ואם כן, באיזו מידה הכללה זו נכונה? מדדי הקשר שנדון בהם כעת וישר הרגרסיה שיבוא אחר כך, עוסקים בשאלות אלו.

3.1 שונות משותפת (Covariance)

הגדרה: נניח שנתונים שני משתנים X,Y. השונות המשותפת להם מוגדרת להיות:

$$\mathrm{Cov}\left(X,Y\right)=\frac{1}{n}\left[\left(X_1-\overline{X}\right)\left(Y_1-\overline{Y}\right)+...+\left(X_n-\overline{X}\right)\left(Y_n-\overline{Y}\right)\right]$$

$$=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)$$

דוגמה: נסדר את הנתונים בטבלה:

X	**Y**	$X_i-\overline{X}$	$Y_i-\overline{Y}$	$\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)$
1	5	-3	-1	3
3	4	-1	-2	2
4	7	0	1	0
5	1	1	-5	-5
7	13	3	7	21

ניתן לראות כי $\overline{X}=4$, $\overline{Y}=6$, $\mathrm{SD}\left(X\right)=2$, $\mathrm{SD}\left(Y\right)=4$. נחשב את השונות המשותפת וניווכח ש-$\mathrm{Cov}\left(X,Y\right)=4.2>0$. המשמעות של העובדה שהשונות המשותפת של X ו-Y חיובית, היא שהמשתנים הללו תלויים באופן חיובי. כלומר, אך בהכללה, אם אחד גדל - גם האחר גדל.

נוסחה: ניתן לראות שמתקיים:

$$\mathrm{Cov}\left(X,Y\right)=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)$$

$$=\frac{1}{n}\sum_{i=1}^{n}X_iY_i-\frac{1}{n}\sum_{i=1}^{n}X_i\overline{Y}-\frac{1}{n}\sum_{i=1}^{n}\overline{X}Y_i+\frac{1}{n}\sum_{i=1}^{n}\overline{X}\cdot\overline{Y}$$

$$=\tfrac{1}{n}\textstyle\sum_{i=1}^{n}X_i\cdot Y_i-\overline{X}\cdot\overline{Y}-\overline{Y}\cdot\overline{X}+\overline{X}\cdot\overline{Y}=\overline{X\cdot Y}-\overline{X}\cdot\overline{Y}$$

ולכן קיבלנו את הנוסחה:

$$\mathrm{Cov}\left(X,Y\right)=\overline{X\cdot Y}-\overline{X}\cdot\overline{Y}$$

מתאם בין משתנים: נאמר שהמשתנים X,Y מתואמים חיובית (שלילית) אם השונות המשותפת שלהם חיובית (שלילית). נאמר שהמשתנים X,Y בלתי־מתואמים אם השונות המשותפת שלהם היא 0.

תכונות השונות המשותפת

1. הזזה באמצעות חיבור קבועים לשני המשתנים (גם אם התזוזות שונות זו מזו בערכן) אינה משנה את השונות המשותפת. כלומר, לכל a, b קבועים מתקיים:

$$\text{Cov}(X+a, Y+b) = \text{Cov}(X, Y)$$

הסיבה לכך היא ששינוי כל התצפיות בקבוע מזיז את הממוצע בדיוק באותו קבוע, ולכן ההפרשים מהממוצע לא משתנים.

2. לכל a, b קבועים מתקיים:

$$\text{Cov}(aX, bY) = ab \cdot \text{Cov}(X, Y)$$

משתי התכונות הללו נובע שלכל a, b, c, d קבועים מתקיים:

$$\text{Cov}(aX+b, cY+d) = ac \cdot \text{Cov}(X, Y)$$

3.

$$\text{Cov}(X, Y) = \text{Cov}(Y, X)$$

4.

$$\text{Cov}(X+Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$$

5.

$$\text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y) \pm 2\text{Cov}(X, Y)$$

הוכחה:

$$\text{Var}(X \pm Y) = \frac{1}{n}\sum_{i=1}^{n}\left[(X_i \pm Y_i) - \left(\overline{X \pm Y}\right)\right]^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left[\left(X_i - \overline{X}\right) \pm \left(Y_i - \overline{Y}\right)\right]^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left[\left(X_i - \overline{X}\right)^2 + \left(Y_i - \overline{Y}\right)^2 \pm 2\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)\right]^2$$

$$= \text{Var}(X) + \text{Var}(Y) \pm 2\text{Cov}(X, Y)$$

3.2 מקדם המתאם (Correlation coefficient)

הגדרה: מקדם המתאם בין X,Y מוגדר להיות:

$$\mathrm{Corr}\left(X,Y\right)=\frac{\mathrm{Cov}\left(X,Y\right)}{\mathrm{SD}\left(X\right)\cdot\mathrm{SD}\left(Y\right)}$$

הרחבה: נוכל לפתח את הביטוי ולקבל:

$$\mathrm{Corr}\left(X,Y\right)=\frac{\mathrm{Cov}\left(X,Y\right)}{\mathrm{SD}\left(X\right)\cdot\mathrm{SD}\left(Y\right)}$$

$$=\frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)}{\mathrm{SD}\left(X\right)\mathrm{SD}\left(Y\right)}=\frac{1}{n}\sum_{i=1}^{n}\frac{\left(X_i-\overline{X}\right)}{\mathrm{SD}\left(X\right)}\cdot\frac{\left(Y_i-\overline{Y}\right)}{\mathrm{SD}\left(Y\right)}$$

נשים לב שקיבלנו שההגדרה של מקדם המתאם של X,Y שקולה לשונות המשותפת של ציוני התקן של X ושל Y.

תכונות מקדם המתאם

1.

$$\mathrm{Corr}\left(X,Y\right)=\mathrm{Corr}\left(Y,X\right)$$

2. לכל a,b,c,d קבועים מתקיים:

$$\mathrm{Corr}\left(aX+b,cY+d\right)=\frac{ac\mathrm{Cov}\left(X,Y\right)}{\left|a\right|\mathrm{SD}\left(X\right)\left|c\right|\mathrm{SD}\left(Y\right)}=\begin{cases}\mathrm{Corr}\left(X,Y\right) & \text{if } a\cdot c>0\\ -\mathrm{Corr}\left(X,Y\right) & \text{if } a\cdot c<0\end{cases}$$

ובפרט מתקיים:

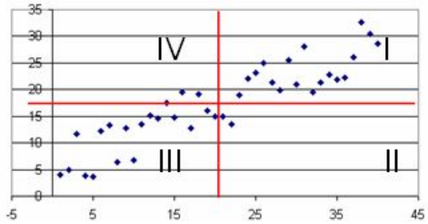
$$\mathrm{Corr}\left(\frac{X-\overline{X}}{SD\left(X\right)},\frac{Y-\overline{Y}}{SD\left(Y\right)}\right)=\mathrm{Corr}\left(X,Y\right)$$

תכונה יסודית של מקדם המתאם:

$$-1\leq\mathrm{Corr}\left(X,Y\right)\leq1$$

(נוכיח טענה זו בהמשך (עמוד 126)).
מתכונה זו נובע שכאשר מקדם המתאם הוא למשל 0.8 מדובר בקשר חזק בין שני המשתנים.

ניתן לנסות לנחש את סימן מקדם המתאם בעזרת התבוננות בדיאגרמת פיזור (ראה הרחבה להלן). נסתכל למשל בציור הבא:



נבחר למקם את ראשית הצירים בנקודת הממוצעים, כלומר $(\overline{X},\overline{Y})$ ונעביר קווים ישרים מקבילים לצירים לאורך ולרוחב דרך נקודה זו. נתחיל ברביע הראשון, באזור זה ערך ה־x וה־y של כל הנקודות גדול מ־$\overline{X}$ ו־$\overline{Y}$ בהתאמה. לכן, אזור זה יתרום לסכום שבהגדרת השונות מכפלות חיוביות. מכיוון שברביע השלישי ערך ה־x וה־y של כל הנקודות קטן מ־$\overline{X}$ ו־$\overline{Y}$ בהתאמה, אזי נקבל הפרשים שליליים מהממוצע ומכפלות חיוביות. יש מעט נקודות ברביע השני והרביעי, שעל פי הגיון זה יניבו מכפלות שליליות, אבל הן מעטות ולכן ננחש שהשונות המשותפת תהיה חיובית.

הרחבה: דיאגרמת פיזור ניתן לתאר את סדרת הנתונים $\{(X_i,Y_i)\}_{i=1}^{n}$ על רשת צירים דו־ מימדית, בעזרת n נקודות במישור. הנקודה (X_i,Y_i) תהא זו שמרחקה האנכי מציר ה־ y הוא X_i, (ייתכן והמרחק שלילי) ומרחקה מציר ה־x הוא Y_i. המטרה היא לא רק לרשום את ה`עצים` אלא גם לראות את ה`יער` , קרי, לבחון באופן גס ולא מדעי האם קיימים קשרים בין המשתנים.

4 רגרסיה לינארית (Linear regression) או: ישר הריבועים הפחותים

נגדיר קו ישר מהצורה $y=b+ax$, כך ש-b הוא החותך ו-a הוא השיפוע. נשים לב כי יחידות bהן ביחידות y ויחידות a הן ביחידות yחלקי x. נרצה לעסוק במרחק שבין סדרה כלשהי של נקודות $\{(X_i,Y_i)\}_{i=1}^{n}$ במישור לבין ישר זה. נגדיר מרחק זה להיות:

$$\sum_{i=1}^{n}(Y_i-(b+aX_i))^2=\sum_{i=1}^{n}(Y_i-b-aX_i)^2$$

נשים לב שכאשר $a=0$, $b=\overline{Y}$ מקבלים את הישר הקבוע $y=\overline{Y}$, ולעיל כשעסקנו בממוצע הגדרנו ביטוי למרחק של הסדרה ממנו:

$$\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2=n\cdot\mathrm{Var}(Y)$$

כאן יש לציין כי אין סמטריה בין X לבין Y. בפרט, המרחק כאן בין נקודה לישר הוא הפרש בין Y_i לבין aX_i+b. ניזכר כי לעיתים (שיעורי פיזיקה) מוגדר אחרת המרחק בין נקודה לישר זהו המרחק בקו אוירי בין (X_i,Y_i) לנקודה הקרובה ביותר אליה על הישר.

כאן קיימת סמטריה בין X ל־Y. כעת נרצה למצוא a, b כאלה שימזערו את המרחק של סדרת הנקודות $\{(X_i, Y_i)\}_{i=1}^n$ מהישר $y = b + ax$. הישר שיתקבל לאחר הצבת הנקודות a, b שנבחר ייקרא **ישר הרגרסיה** או **ישר הריבועים הפחותים** של Y על X. בישר זה נשתמש כדי **לחזות** את Y בהינתן X. למשל בהינתן X_3 נצפה ש־Y_3 יהיה שווה ל־$b + aX_3$. פעמים רבות תחזית זו תהיה לא נכונה, וייווצר פער בין Y_3 האמיתי לבין Y_3 החזוי שמכונה **שארית**. המטרה היא למזער את השאריות ככל שניתן.

הערה: בניגוד לשונות המשותפת ולמקדם המתאם בהם יש סימטריה ביחס שבין X ל־Y, בישר הרגרסיה הדבר לא־כך, והישר של X על Y שונה מהישר של Y על X.

משפט: ישר הרגרסיה של Y על X הוא הישר (היחיד) שעובר דרך הנקודה $\left(\overline{X}, \overline{Y}\right)$ וששיפועו הוא $a = \text{Corr}\left(X, Y\right) \cdot \frac{SD(Y)}{SD(X)}$. או באופן שקול: הישר (היחיד) שהחותך שלו הוא $b = \overline{Y} - a\overline{X}$ וששיפועו הוא:

$$a = \text{Corr}\left(X, Y\right) \cdot \frac{\text{SD}\left(Y\right)}{\text{SD}\left(X\right)}$$

הערה: נשים לב שמתקיים:

$$a = \text{Corr}\left(X, Y\right) \cdot \frac{\text{SD}\left(Y\right)}{\text{SD}\left(X\right)} = \frac{\text{Cov}\left(X, Y\right)}{\text{SD}\left(X\right)\text{SD}\left(Y\right)} \cdot \frac{\text{SD}\left(Y\right)}{\text{SD}\left(X\right)} = \frac{\text{Cov}\left(X, Y\right)}{\text{Var}\left(X\right)}$$

הוכחה: ראשית נשים לב שבהינתן a קבוע כלשהו, הערך של b שממזער את הביטוי $\sum_{i=1}^n \left(Y_i - b - aX_i\right)^2$ הוא $b = \overline{Y} - a\overline{X}$. טענה זו נובעת מכך שהראינו לעיל שמיזעור פונקציית הפסד מסוג זה מתקבל באמצעות הממוצע של הסדרה שבמקרה שלנו היא $\{Y_i - aX_i\}_{i=1}^n$. ולכן בהינתן a נבחר את b להיות $b = \overline{Y - aX} = \overline{Y} - a\overline{X}$. נשים לב שטענה זו מספיקה כדי להראות שהנקודה $\left(\overline{X}, \overline{Y}\right)$ על הישר המבוקש, שכן היא מקיימת את המשוואה $\overline{Y} = b + a\overline{X}$ עבור b שמצאנו לכל ערך של a.

אם כן נותר למצוא a שימזער את הביטוי $\sum_{i=1}^n \left(Y_i - \left(\overline{Y} - a\overline{X}\right) - aX_i\right)^2$.
נפתח את הביטוי באופן הבא:

$$\sum_{i=1}^n \left(Y_i - \left(\overline{Y} - a\overline{X}\right) - aX_i\right)^2 = \sum_{i=1}^n \left[\left(Y_i - \overline{Y}\right) - a\left(X_i - \overline{X}\right)\right]^2$$

$$= \sum_{i=1}^n \left(Y_i - \overline{Y}\right)^2 - 2a\sum_{i=1}^n \left(Y_i - \overline{Y}\right)\left(X_i - \overline{X}\right) + a^2\sum_{i=1}^n \left(X_i - \overline{X}\right)^2$$

כעת נשים לב שניתן להתייחס לביטוי שהתקבל כאל פרבולה צוחקת כשהמשתנה הוא a.
לפרבולה צוחקת מהצורה $cx^2 + dx + e$ יש מינימום שמתקבל על־ידי הנוסחה $x = \frac{-d}{2c}$, ואם נציב במקרה שלנו נקבל שהמינימום מתקבל:

$$\frac{2\sum_{i=1}^n \left(Y_i - \overline{Y}\right)\left(X_i - \overline{X}\right)}{2\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} = \frac{\frac{1}{n}\sum_{i=1}^n \left(Y_i - \overline{Y}\right)\left(X_i - \overline{X}\right)}{\frac{1}{n}\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} = \frac{\text{Cov}\left(X, Y\right)}{\text{Var}\left(X\right)}$$

■

הרחבה: עבור $y=cx^2+dx+e$, $c>0$ (פרבולה צוחקת) הערך המינימלי מתקבל באמצעות הביטוי $y=e-\frac{d^2}{4c}$, נציב את הערכים של הביטוי שקיבלנו:

$$\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2-\frac{\left[-2\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)\left(X_i-\overline{X}\right)\right]^2}{4\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}$$

$$=\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2-\frac{\left[\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)\left(X_i-\overline{X}\right)\right]^2}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}$$

$$=\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2\left[1-\frac{\left[\frac{1}{n}\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)\left(X_i-\overline{X}\right)\right]^2}{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2\cdot\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}\right]$$

$$=\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2\left[1-\mathrm{Corr}^2\left(X,Y\right)\right]=nVar(Y)[1-\mathrm{Corr}^2\left(X,Y\right)]$$

ראשית ניכר שככל שהביטוי $\mathrm{Corr}^2\left(X,Y\right)$ קרוב יותר ל-1, סכום ריבועי השאריות מישר הרגרסיה של Y על X קטן יחסית לסכום ריבועי השאריות מ-$\overline{Y}$. כעת ניזכר בכך שהביטוי כולו התקבל כסכום של ריבועים ולכן הוא לא יכול להיות שלילי, ומכאן נסיק כי $-1\le\mathrm{Corr}\left(X,Y\right)\le 1$, כפי שטענו לעיל מבלי להוכיח.

מסקנה: נשים לב שהמשמעות של מקרה בו השאריות מתאפסות, היא שכל הנקודות $\{(X_i,Y_i)\}_{i=1}^{n}$ ממוקמות על ישר אחד. נזכור שהביטוי שמתאר את השאריות הוא:

$$\sum_{i=1}^{n}\left(Y_i-b-aX_i\right)^2=\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2\left[1-\mathrm{Corr}^2\left(X,Y\right)\right]=0$$

ומכאן שהתאפסות מתרחשת אם ורק אם $\mathrm{Corr}^2\left(X,Y\right)=1$. נסיק שכל הנקודות ממוקמות על ישר אחד אם ורק אם $\mathrm{Corr}^2\left(X,Y\right)=1$.

סימונים:

$$S_{xx}=\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2=nVar(X)$$

$$S_{yy}=\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2=nVar(Y)$$

$$S_{xy}=\sum_{i=1}^{n}\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)=nCorr(X,Y)$$

$$S_{ee}=\sum_{i=1}^{n}\left(Y_i-\hat{Y}_i\right)^2, \qquad \hat{y}=b+ax \quad , \quad \hat{Y}_i=b+aX_i \;\; ,1\le i\le n$$

$$S_{\hat{y}\hat{y}}=\sum_{i=1}^{n}\left(\hat{Y}_i-\overline{Y}\right)^2$$

$\hat{Y}_i$ יקרא הערך החזוי ה־i בעזרת ישר הרגרסיה
נשים לב שבסימונים אלה, שיפוע ישר הרגרסיה הוא $a=\frac{S_{xy}}{S_{xx}}$.
נסמן את מידת הקירבה של הנקודות לישר הרגרסיה בהשוואה למידת הקירבה של הנקודות לממוצע:

$$\mathrm{Corr}^2\left(X,Y\right)=R^2$$

וכן נשים לב שבסימונים אלה מתקיים:

$$1-R^2=\frac{S_{ee}}{S_{yy}}$$

תכונות ישר הרגרסיה

1. הישר עובר דרך הנקודה $\left(\overline{X},\overline{Y}\right)$.
2. נסמן את סדרת השאריות $\left\{Y_i-\hat{Y}_i\right\}_{i=1}^{n}=\{e_i\}_{i=1}^{n}$.
   מתקיים $\sum_{i=1}^{n}e_i=0$, או באופן שקול $\bar{e}=0$.
   <u>הוכחה:</u>

$$\sum_{i=1}^{n}e_i=\sum_{i=1}^{n}\left(Y_i-\hat{Y}_i\right)=\sum_{i=1}^{n}\left(Y_i-b-aX_i\right)=\sum_{i=1}^{n}Y_i-\sum_{i=1}^{n}b-\sum_{i=1}^{n}aX_i=$$

$$=n\overline{Y}-nb-an\overline{X}=n\overline{Y}-n\left(\overline{Y}-a\overline{X}\right)-an\overline{X}=0.$$

3. ממוצע הערכים החזויים שווה לממוצע של הערכים הנכונים. כלומר:

$$\overline{\hat{Y}} = \overline{Y}$$

<u>הוכחה:</u>

$$\overline{\hat{Y}} = \overline{aX+b} = a\overline{X} + b = \overline{Y}$$

<u>הערה:</u> תכונות $1-3$ תקפות לכל ישר העובר בנקודה $\left(\overline{X}, \overline{Y}\right)$ ולא רק לישר הרגרסיה.

4. מתקיים $\sum_{i=1}^{n} X_i e_i = 0$, או, באופן שקול, $\overline{Xe} = 0$, ז`א ממוצע משוקלל של השגיאות כאשר כל שגיאה מקבלת משקל יחסי לערך X המתאים לה (היכול להיות שלילי), הוא אפס.
<u>הוכחה:</u>

$$\sum_{i=1}^{n} X_i e_i = \sum_{i=1}^{n} X_i \left(Y_i - \hat{Y}_i\right) = \sum_{i=1}^{n} X_i \left(Y_i - b - aX_i\right)$$

$$= \sum_{i=1}^{n} X_i \left(Y_i - \overline{Y} + a\overline{X} - aX_i\right) = \sum_{i=1}^{n} X_i \left(Y_i - \overline{Y}\right) - a\sum_{i=1}^{n} X_i \left(X_i - \overline{X}\right)$$

$$= \sum_{i=1}^{n} \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right) + \overline{X}\underbrace{\sum_{i=1}^{n} \left(Y_i - \overline{Y}\right)}_{=0} - a\sum_{i=1}^{n} \left(X_i - \overline{X}\right)\left(X_i - \overline{X}\right) + a\overline{X}\underbrace{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)}_{=0}$$

$$= \sum_{i=1}^{n} \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right) - a\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$$

$$= n\mathrm{Cov}\left(X, Y\right) - \frac{\mathrm{Cov}\left(X, Y\right)}{\mathrm{Var}\left(X\right)} n\mathrm{Var}\left(X\right) = n\mathrm{Cov}\left(X, Y\right) - n\mathrm{Cov}\left(X, Y\right) = 0$$

5. מתקיים $\sum_{i=1}^{n} \hat{Y}_i e_i = 0$, או באופן שקול $\overline{\hat{Y}e} = 0$.
<u>הוכחה:</u>

$$\sum_{i=1}^{n} \hat{Y}_i e_i = \sum_{i=1}^{n} \left(b + aX_i\right) e_i = b\sum_{i=1}^{n} e_i + a\sum_{i=1}^{n} X_i e_i = b \cdot 0 + a \cdot 0 = 0$$

כמו ב–4 רק משקולות $\hat{Y}_i$.

6.

$$\sum_{i=1}^{n} \left(Y_i - \overline{Y}\right)^2 = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2 + \sum_{i=1}^{n} \left(\hat{Y}_i - \overline{Y}\right)^2$$

או ברישום מקוצר:

$$S_{yy} = S_{ee} + S_{\hat{y}\hat{y}}$$

<u>הוכחה:</u>

$$\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2=\sum_{i=1}^{n}\left(Y_i-\hat{Y}_i+\hat{Y}_i-\overline{Y}\right)^2$$

$$=\sum_{i=1}^{n}\left[\left(Y_i-\hat{Y}_i\right)^2+2\left(Y_i-\hat{Y}_i\right)\left(\hat{Y}_i-\overline{Y}\right)+\left(\hat{Y}_i-\overline{Y}\right)^2\right]$$

$$=\sum_{i=1}^{n}\left(Y_i-\hat{Y}_i\right)^2+2\sum_{i=1}^{n}\left(Y_i-\hat{Y}_i\right)\left(\hat{Y}_i-\overline{Y}\right)+\sum_{i=1}^{n}\left(\hat{Y}_i-\overline{Y}\right)^2$$

$$=\sum_{i=1}^{n}\left(Y_i-\hat{Y}_i\right)^2+\sum_{i=1}^{n}\left(\hat{Y}_i-\overline{Y}\right)^2$$

השוויון האחרון נכון מכיוון שמתקיים:

$$\sum_{i=1}^{n}\left(Y_i-\hat{Y}_i\right)\left(\hat{Y}_i-\overline{Y}\right)=\sum_{i=1}^{n}e_i\hat{Y}_i-\sum_{i=1}^{n}e_i\overline{Y}=0-\overline{Y}\sum_{i=1}^{n}e_i=0$$

7. <u>מסקנה:</u>

$$\mathrm{Var}\left(Y\right)=\frac{1}{n}\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2=\frac{1}{n}\sum_{i=1}^{n}\left(Y_i-\hat{Y}_i\right)^2+\frac{1}{n}\sum_{i=1}^{n}\left(\hat{Y}_i-\overline{Y}\right)^2$$

הגודל $\left(\hat{Y}_i-\overline{Y}\right)^2$ $\frac{1}{n}\sum_{i=1}^{n}$ נקרא `השונות המוסברת` (של Y באמצעות ישר הרגרסיה של Y על X) . אכן, ככל שהוא גדול יותר, זאת אומרת, קרוב יותר (מלמטה) לשונות של Y, אנו חשים כי מרבית הפיזור סביב $\overline{Y}$ של Y באה לידי ביטוי בעזרת ישר הרגרסיה. מן הסתם, $\frac{1}{n}\sum_{i=1}^{n}\left(Y_i-\hat{Y}_i\right)^2$ היא `השונות הלא מוסברת`

8. ביטוי אלטרנטיבי לישר הרגרסיה:

$$\frac{y-\overline{Y}}{\mathrm{SD}\left(Y\right)}=\mathrm{Corr}\left(X,Y\right)\cdot\frac{x-\overline{X}}{\mathrm{SD}\left(X\right)}$$

9. ריבוע מקדם המתאם מודד עד כמה צפופות הנקודות סביב ישר הרגרסיה, בהשוואה לצפיפותן סביב הישר (הקבוע) $y=\overline{Y}$.

10.

$$\mathrm{Var}\left(\hat{Y}\right)=\mathrm{Corr}r^2\left(X,Y\right)\mathrm{Var}\left(Y\right)$$

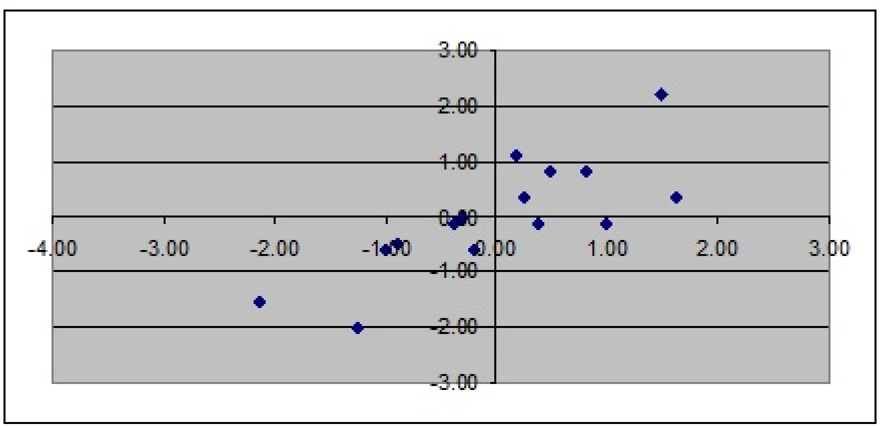
11. הסימן של $\mathrm{Corr}\left(X,Y\right)$ מאפיין את היחס בין המשתנה המתוקנן של Y_i לבין המשתנה המתוקנן של X_i: למשל, אם $Corr(X,Y) > 0$. ז"א קיים מתאם חיובי בין X לבין Y אז, בהכללה, ניתן לומר כי ככל ש X גדל כך גדל גם Y.

12. נניח כי שיפוע ישר הרגרסיה של Y על X חיובי. נבצע רוטציה בזווית מסוימת של הנתונים כך שהם יעברו תנועה מעגלית עם כיוון השעון, כאשר הציר בנקודה $\left(\overline{X},\overline{Y}\right)$, ונדאג רק שהקורלציה תשאר חיובית גם לאחר הרוטציה. השיפוע של ישר הרגרסיה החדש יקטן בהשוואה לקודמו. הממוצעים של X ושל Y ישתנו מעט. $\mathrm{SD}\left(X\right)$ עלה במעט אך $\mathrm{SD}\left(Y\right)$ קטן באופן משמעותי. לבסוף, S_{ee} ישתנה במעט. מכאן ש-$\frac{S_{ee}}{S_{yy}}$ יגדל באופן משמעותי. נשים לב כי למרות שצפיפות הנקודות סביב ישר הרגרסיה נותרה כמעט ללא שינוי, מקדם המתאם ירד משמעותית. ההסבר לכך הוא שפיזור הנקודות סביב הישר $y = \overline{Y}$ קטן משמעותית, ולכן הצפיפות של הנקודות סביב ישר הרגרסיה בהשוואה לצפיפות סביב הישר הנ`ל - קטנה.
<u>מסקנה</u>: ככל ש-$\mathrm{Corr}^2\left(X,Y\right)$ גדול יותר כך פיזור הנקודות סביב ישר הרגרסיה בהשוואה לפיזורן סביב ישר הממוצע - נמוך יותר. (פיזור במובן של סכום ריבועי הסטיות).
<u>לסיכום</u>: ככל שהמתאם גדול יותר קו הרגרסיה של Y על X מנבא טוב יותר בהשוואה לממוצע.

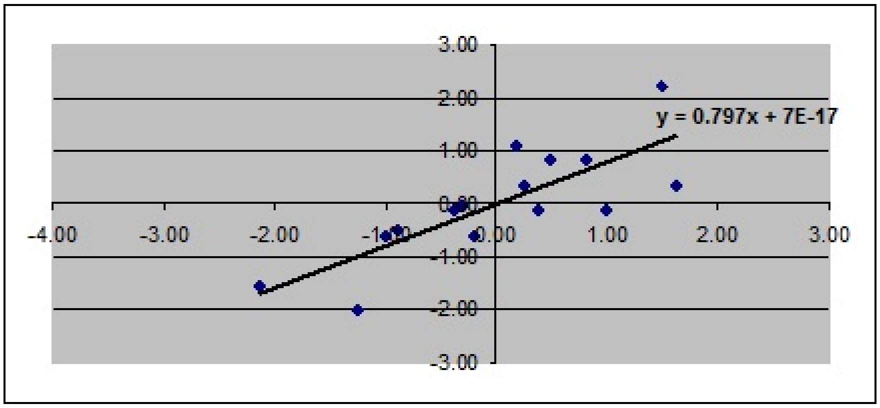
דוגמה: נתבונן בנתונים הבאים:

−2.12	0.80	0.24	−0.19	−0.32	0.49	0.36	−0.88	−1.00	0.99	0.18	−0.38	−1.25	1.48	1.61	**X**
−1.53	0.81	0.34	−0.59	−0.03	0.81	−0.12	−0.50	−0.59	−0.12	1.09	−0.12	−2.00	2.22	0.34	**Y**

נחשב ונקבל $\overline{X} = \overline{Y} \simeq 0$, וכן $\sqrt{R^2} = \sqrt{\mathrm{Corr}^2\left(X,Y\right)} = 0.635$.
דיאגרמת הפיזור של הנתונים במערכת צירים היא:



נתאים ישר רגרסיה בהתאם לנוסחה שהוכחנו, ונקבל את הישר הבא:



כעת ננתח שינוי שבו מבצעים רוטציה לנתונים, ומביאים להטיה בזווית כלשהי את הישר עם כיוון השעון, סביב ראשית הצירים $\left(\overline{X},\overline{Y}\right)=(0,0)$.
נגדיר כעת סדרה של תצפיות חדשות $\left(X^{*},Y^{*}\right)$, ונניח לדוגמה שכל אחת מהתצפיות מתקבלת מהסדרה המקורית באמצעות הביטוי:

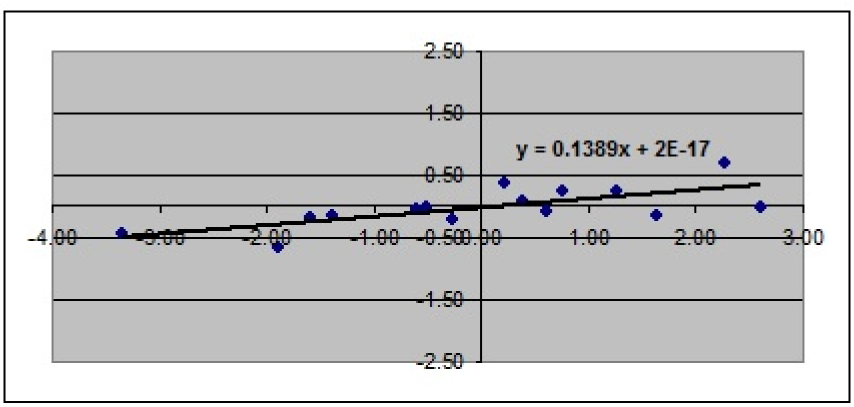
$$x^{*}=1.63x-0.07y$$

$$y^{*}=-0.07x+0.37y$$

נקבל במקרה זה את סדרת התצפיות:

-3.35	1.25	0.37	-0.28	-0.52	0.74	0.60	-1.40	-1.59	1.62	0.21	-0.61	-1.90	2.26	2.60	$\mathbf{X}^{*}$
-0.42	0.24	0.11	-0.21	0.01	0.27	-0.07	-0.12	-0.15	-0.12	0.39	-0.02	-0.65	0.72	0.01	$\mathbf{Y}^{*}$

נחשב ונקבל $\overline{X}=\overline{Y}\simeq 0$, וכן $\sqrt{R^{2}}=\sqrt{\mathrm{Corr}^{2}\left(X,Y\right)}=0.48$.
דיאגרמת הפיזור של הנתונים החדשים בתוספת ישר הרגרסיה החדש, היא:



נשים לב שישר הרגרסיה החדש עדיין עובר ב־$\left(\overline{X},\overline{Y}\right)=(0,0)$, אבל הוא קרוב יותר לישר הקבוע $y=\overline{Y}=0$, כי שיפועו קטן (אך עם זאת נשאר בעל אותו סימן). כמו־כן נשים לב

שגם מקדם המתאם ירד באופן משמעותי, מ־0.8 ל־0.7.[4] כלומר, לאחר ביצוע הרוטציה ירד טיב הניבוי של ישר הרגרסיה.

4.1 נסיגה לממוצע (Regression to the mean)

ראינו שישר הרגרסיה מתקבל מהמשוואה:

$$\frac{y-\overline{Y}}{\mathrm{SD}\left(Y\right)}=\mathrm{Corr}\left(X,Y\right)\frac{x-\overline{X}}{\mathrm{SD}\left(X\right)}$$

נניח שאחד הנתונים X_i נמצא k סטיות תקן מעל (מתחת) לממוצע $\overline{X}$, אז ציון התקן $Z_i=\frac{X_i-\overline{X}}{\mathrm{SD}(X)}$ שווה k ($-k$). לפי הניבוי של ישר הרגרסיה, הנתון $\hat{Y}_i$ יהיה במרחק של $k\cdot\mathrm{Corr}\left(X,Y\right)$ סטיות תקן מעל (מתחת) לממוצע $\overline{Y}$. נשים לב שהניבוי $k\cdot\mathrm{Corr}\left(X,Y\right)$ יתקבל תמיד בטווח $\left[\overline{Y}-k\cdot\mathrm{SD}\left(Y\right),\overline{Y}+k\cdot\mathrm{SD}\left(Y\right)\right]$.

הנסיגה לממוצע קובעת שבמונחי ציון תקן, המרחק של $\hat{Y}$ (הערך החזוי) מהממוצע $\overline{Y}$ קטן מהמרחק של X מהממוצע $\overline{X}$. עובדה זו נובעת מכך שהקורלציה $\mathrm{Corr}\left(X,Y\right)$ תמיד קטנה בערכה המוחלט מ־1.

ישר סטיות התקן: נגדיר את ישר סטיות התקן להיות $\frac{y-\overline{Y}}{\mathrm{SD}(Y)}=\frac{x-\overline{X}}{\mathrm{SD}(X)}$. או באופן שקול לאחר העברת אגפים:

$$y=\frac{\mathrm{SD}\left(Y\right)}{\mathrm{SD}\left(X\right)}\left(x-\overline{X}\right)+\overline{Y}=\frac{\mathrm{SD}\left(Y\right)}{\mathrm{SD}\left(X\right)}x-\frac{\mathrm{SD}\left(Y\right)}{\mathrm{SD}\left(X\right)}\overline{X}+\overline{Y}$$

ישר סטיות התקן אינו מאופיין בנסיגה אל הממוצע והוא מגדיר מתאם מלא. כלומר, בתחזית הנקבעת לפי ישר סטיות התקן, אם X_i נמצא k סטיות תקן מעל (מתחת) $\overline{X}$, אז גם $\hat{Y}$ נמצא k סטיות תקן מעל (מתחת) $\overline{Y}$. נשים לב שישר הרגרסיה הוא הישר שממזער את השאריות כך שישר סטיות התקן בהכרח פחות טוב ממנו.

הערה חשובה: קורלציה אינה זהה לסיבתיות (causality)!
כלומר, העובדה שקיים מתאם בין שני משתנים אינה אומרת ששינוי באחד יוביל לשינוי באחר. כך למשל קיימת קורלציה חזקה בין משקל לבין גובה, ועם זאת השמנה אינה גוררת עלייה בגובה.

ישר הרגרסיה של X על Y

1. ישר הרגרסיה של X על Y שונה מישר הרגרסיה של Y על X.

2. נזכור שישר הרגרסיה של Y על X הוא:

$$\frac{y-\overline{Y}}{\mathrm{SD}\left(Y\right)}=\mathrm{Corr}\left(X,Y\right)\frac{\left(x-\overline{X}\right)}{\mathrm{SD}\left(X\right)}$$

[4] קיבלנו את המספרים האלה מכך ש: $0.7\simeq\sqrt{0.48}$, $0.8\simeq\sqrt{0.635}$.

ולכן אם נחליף תפקידים ונזכור כי $Corr\left(X,Y\right)=Corr\left(Y,X\right)$, נקבל את ישר הרגרסיה של X על Y:

$$\frac{x-\overline{X}}{\mathrm{SD}\left(X\right)}=\mathrm{Corr}\left(X,Y\right)\frac{\left(y-\overline{Y}\right)}{\mathrm{SD}\left(Y\right)}$$

ולכן הישר על אותה מערכת צירים ייראה מהצורה:

$$y-\overline{Y}=\frac{\mathrm{SD}\left(Y\right)}{\mathrm{SD}\left(X\right)}\frac{\left(x-\overline{X}\right)}{\mathrm{Corr}\left(X,Y\right)}$$

3. באותה מערכת צירים, שני ישרי הרגרסיה הללו נחתכים בנקודה $\left(\overline{X},\overline{Y}\right)$, וכן שיפועו של ישר הרגרסיה של X על Y ($\frac{\mathrm{SD}(Y)}{\mathrm{SD}(X)}\frac{1}{\mathrm{Corr}(X,Y)}$) חד יותר משיפוע ישר הרגרסיה של Y על X ($\mathrm{Corr}\left(X,Y\right)\frac{\mathrm{SD}(Y)}{\mathrm{SD}(X)}$). הסיבה: $-1<Corr(X,Y)<1$
4. השיפוע של ישר סטיות התקן הוא ערך ביניים כלשהו בין השיפועים הללו. לא נוכיח, אולם $\mathrm{SD}\left(X\right)=\mathrm{SD}\left(Y\right)$ אם ורק אם ישר סטיות התקן הוא חוצה הזווית שבין שני ישרי הרגרסיה הנ"ל.
5. ישר סטיות התקן של X על Y מתלכד עם ישר סטיות התקן של Y על X.
6. הראינו לעיל כי עבור ישר הרגרסיה של Y על X מתקיים:

$$\frac{S_{ee}}{S_{yy}}=1-\mathrm{Corr}^{2}\left(X,Y\right)$$

כאשר $S_{ee}=\sum_{i=1}^{n}\left(Y_i-\hat{Y}_i\right)^2$ וכן $\hat{Y}_i=\mathrm{Corr}\left(X,Y\right)\frac{\mathrm{SD}(Y)}{\mathrm{SD}(X)}\left(X_i-\overline{X}\right)+\overline{Y}$.
באותו אופן עבור ישר הרגרסיה של X על Y מתקיים:

$$\frac{S_{ee}^{*}}{S_{xx}}=1-\mathrm{Corr}^{2}\left(X,Y\right)$$

כאשר $S_{ee}^{*}=\sum_{i=1}^{n}\left(X_i-\hat{X}_i\right)$ וכן $\hat{X}_i=\mathrm{Corr}\left(X,Y\right)\frac{\mathrm{SD}(X)}{\mathrm{SD}(Y)}\left(Y_i-\overline{Y}\right)+\overline{X}$.
נסיק מכך:

$$\frac{S_{ee}^{*}}{S_{ee}}=\frac{S_{xx}}{S_{yy}}=\frac{\mathrm{Var}\left(X\right)}{\mathrm{Var}\left(Y\right)}$$

במילים, בין שני ישרי הרגרסיה יחס סכום ריבועי הסטיות ששווה ליחס בין השונויות. כמו־כן, מכיוון ש-$\mathrm{Corr}\left(X,Y\right)=\mathrm{Corr}\left(Y,X\right)$, שיעור השונות המוסברת שווה בשני ישרי הרגרסיה. כלומר, יכולת הניבוי של שני ישרי הרגרסיה שווה בעוצמתה.

5 אמידה

5.1 ממוצע הממוצעים וסטיית התקן של הממוצעים

נניח כי האוכלוסיה מורכבת מ־n תצפיות $\{\mathrm{x}_i\}_{i=1}^{n}$ עם ממוצע $\bar{X}$ ושונות $\mathrm{Var}\left(X\right)$. לצורך הנוחות נניח כי כל המספרים באוכלוסיה שונים זה מזה. נרכיב כעת את כל הסדרות

באורך של m מתוכם. נשים לב כי אנו מרשים חזרה מספר פעמים על אותו איבר ואיננו מגבילים את גודלו של m ביחס ל־n, בפרט ייתכן כי m> n. ניתן לראות כי נקבל $\mathrm{n^m}$ סדרות שונות שכאלו. אם נחשב את ממוצע m האיברים שנרשמו בכל אחת מהסדרות, נקבל סדרה חדשה באורך $\mathrm{n^m}$. כמובן לא כל איברי הסדרה החדשה שונים אלו מאלו. נסמן סדרה זו ב־$\{\overline{X}_{(m)}\}$. נטען בשלב זה ללא הוכחה כי $\overline{\overline{X}}_{(m)} = \overline{X}$ וכן כי $\mathrm{Var}\big(\bar{X}_{(\mathrm{m})}\big) = \frac{\mathrm{Var}(X)}{m}$. כלומר, ממוצע הממוצעים מתלכד עם הממוצע באוכלוסיה והפיזור בממוצעים אלו, הנמדד בסטיית התקן, קטן בפקטור של $\sqrt{\mathrm{m}}$ מהפיזור באוכלוסיה.

ניתן על אוכלוסית הממוצעים להגדיר קבוצות. למשל, נגדיר את הקבוצה $\mathrm{A_x}$ להיות כל הממוצעים אשר ערכם קטן או שווה ל־x. חלק מהממוצעים מקיים תנאי זה. השאר אינם מקיימים. לקבוצה A יש שכיחות יחסית בסדרת הממוצעים ונסמנה ב־$f_m(\mathrm{A})$. רישום זה בא להדגיש כי השכיחות היחסית היא פונקציה של מספר איברים באוכלוסיה עליהם מחושב הממוצע. בפרט, $f_m(\mathrm{A})$ זו השכיחות היחסית של הממוצעים מתוך $\mathrm{n^m}$ הממוצעים שמקורם ב־ m פריטים (לא בהכרח שונים) באוכלוסיה המקיימים את התנאי המאפיין את הקבוצה A. יש לנו עניין מיוחד בקבוצות מהסוג $A_{\overline{X}+\frac{SD(X)}{\sqrt{m}}\zeta}$ כי הם מתייחסים לאותו חלק של סדרת הממוצעים שעבורם ציון התקן קטן מהגודל המסומן, קרי מ ζ.

דוגמא לשכיחות יחסית:

נניח את קבוצת המספרים הבאה: {2,3,5}. ונניח שרוצים לבנות מתוכה סדרה בת 4 איברים. כמה סדרות כאלו קיימות? גודל הקבוצה הוא $n=3$ ואנו מעוניינים בסדרה באורך $m=4$ ולכן קיימות n^m סדרות כאלה. ובמקרה שלנו, 81 סדרות. דוגמאות לסדרה: $\{2,2,3,5\}$, $\{2,2,2,2\}$, $\{3,3,3,5\}$ וכו'.

כעת, נחשב לכל סדרה את ממוצע המספרים בה ונביט בסדרת הממוצעים, $\overline{X}_{(m)}$. מספר האיברים בסדרת הממוצעים הוא 81, כאמור. עבור הדוגמאות לעיל הממוצעים הם $3.5, 2, 3$ בהתאמה.

השאלה שמעניינת אותנו כעת היא, מה השכיחות היחסית של איבר מסויים בסדרה? למשל, מה היא השכיחות היחסית של 3? 2.75? אם כן, עלינו להביט כמה פעמים מופיע הממוצע 3 בסדרה $\overline{X}_{(m)}$ ולחשב את היחס שלו מתוך 81. התוצאה שתתקבל היא השכיחות היחסית של 3. באופן דומה, מעניין אותנו לדעת את השכיחות היחסית של איברי הסדרה הקטנים מערכים מסויימים.

5.2 משפט הגבול המרכזי

משפט: קיימת פונקציה, נסמנה ב־ $\Phi(\zeta)$, כך ש:

$$\lim_{m\to\infty} f_m\left(A_{\overline{X}+\frac{SD(X)}{\sqrt{m}}\zeta}\right) = \Phi(\zeta)$$

יתר על כן,

$$\Phi(\varsigma) = \int_{t=-\infty}^{\varsigma} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

משפט הגבול המרכזי אומר כי לא משנה מהי אוכלוסית המוצא שלנו, ההתנהגות של סדרת הממוצעים, כאשר היא מבוססת על מספר פריטים רב, זהה. בפרט, השכיחות היחסית של אברי סדרת הממוצעים המתוקננים שואפת לגבול אוניברסלי שנגזר מהפונקציה $\Phi(\zeta)$.

הערה: לפונקציה $\Phi(\zeta)$ אין ביטוי מפורש כדי לחשבו יש להשתמש בקירובים נומריים לאינטגרל המגדיר פונקציה זו.

הערה: מאחר ותקנון סדרת הממוצעים הוא שקול לתקנון סכום האיברים המגדירים את הממוצע (לפני החלוקה ב־m) ומאחר ותקנון אינו אלא שינוי סקלה, הטענה על ההתנהגות האוניברסלית המצוינת במשפט הגבול המרכזי, היא למעשה טענה על הסכום. בפרט, סכום של מספר איברים מסדרה מוליך לסדרה אשר התנהגותה אוניברסלית.

נסמן ב־Z_α את המספר כך שהשכיחות היחסית של הערכים הקטנים או שווים לו על פי $\Phi(\varsigma)$ היא α. בפרט, Z_α הוא המספר היחידי המקיים,

$$\int_{t=-\infty}^{Z_\alpha} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \alpha$$

למעשה, Z_α הוא האחוזון ה100α.

הערה: במבט ראשון, התוצאה נראית מוזרה ולא לעניין: בידינו אוכלוסיה סופית ויש לקחת מדגם אינסופי (שודאי יכלול חזרות רבות) על מנת לקבל את השכיחות היחסית הגבולית. אך המציאות שונה. בדרך כלל n הוא הגדול ועבור ערכי m בינוניים למדי ובוודאי קטנים ביחס ל־ n, בכל זאת הקרוב שמושרה על ידי משפט הגבול המרכזי הוא טוב. במילים אחרות, לקחת מדגם בגודל בינוני וההנחה לצורכי עבודה כי השכיחות היחסית על פני אוסף המדגמים שבאים בחשבון היא בעל שכיחויות יחסיות המושרות על ידי הפונקציה $\Phi(\varsigma)$ היא קרובה למציאות. שימוש ב $\Phi(\varsigma)$ כקרוב לפונקציית שכיחות יחסית, נקרא הקירוב הנורמלי. למען הסר ספק נאמר כי אין הטענה מתייחסת לשכיחות היחסית במדגם שנלקח (אלא על פני כל המדגמים הבאים בחשבון).

5.3 אמידה נקודתית

עבור סדרות נתונות יש עניין רב במציאת $\bar{X}$. מאחר ובדרך כלל האוכלוסיות גדולות חישובו המדויק קשה ולעיתים אף בלתי אפשרי. כאן באה לעזרתינו טכניקה הקרויה אמידה שבה

ניתן הערכה לגודל זה. היא בדרך כלל מבוססת על חישוב הממוצע רק על פני מדגם (הקרוי מדגם מקרי). הפונקציה הזו נקראת אומד והתוצאה בפועל הנגזרת מהפעלת הפונקציה על נתוני המדגם נקראת אומדן. ואכן נגדיר ב־ $\bar{X}_{(\mathrm{m})}$ את הממוצע של סדרה **אחת** מתוך n^{m} סדרות הממוצעים שהוגדרו בסעיף הקודם. שגיאתנו באמידת $\bar{X}$ היא כמובן $\bar{X}_{(\mathrm{m})} - \bar{X}$.

שאלה: מהו ממוצע השגיאות על פני כל n^{m} ערכי $\bar{X}_{(\mathrm{m})}$ האפשריים?

תשובה: אפס.

הוכחה: ראינו כי ממוצע סדרת הממוצעים הוא $\bar{X}$. כמו כן, החסרת קבוע מאברי הסדרה, גורמת לשינוי זהה בממוצע (ראו עמוד 7), ולכן הממוצע על פני ערכי $\bar{X}_{(\mathrm{m})} - \bar{X}$ הוא אפס.

שאלה: מהי סטיית התקן של שגיאות אילו?

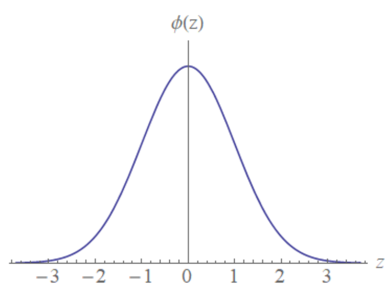
תשובה: $\mathrm{SD(X)}/\sqrt{m}$

הוכחה: ראינו כי סטיית התקן של $\bar{X}_{(\mathrm{m})}$ והזזה בקבוע (במקרה זה, ב־$\bar{X}$) אינה משנה את ערכה (ראו עמוד 17).

אומד לגודל מסויים יקרא חסר הטיה אם ממוצע השגיאות שלו שווה לאפס. אנו רואים כי $\bar{X}_{(\mathrm{m})}$ הוא אומד חסר הטיה ל־$\bar{X}$.

הטבלה הנורמלית:

בנספח הספר (או הפרק) מופיעה הטבלה הנורמלית. נסביר כעת באמצעות דוגמאות כיצד משתמשים בה.

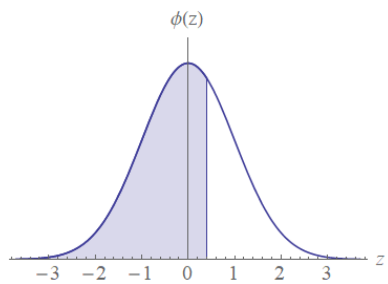


הטבלה הנורמלית מספקת לנו שני נתונים שעשויים לעניין אותנו: ς ו־ $\Phi(\varsigma)$. בהינתן אחד מהערכים הללו, ניתן בעזרת הטבלה למצוא את הערך השני. על מנת להיות אחידים עם הסימונים האוניברסליים, נסמן את השכיחות היחסית ב־ α ואת ς נחליף ב Z.
Z_α יסמן את המספר כך שהשכיחות היחסית של המספרים הקטנים או שווים לו על פי $\Phi(\varsigma)$ היא α. כלומר $\Phi(Z_\alpha) = \alpha$. הפונקציה שהאינטגרל המסוים שלה הוא Φ היא סימטרית סביב אפס, מכך אנו מקבלים את התכונות הבאות:
1) $\Phi(-Z) = 1 - \Phi(Z)$
2) $Z_\alpha = -Z_{1-\alpha}$
<u>מציאת α:</u>
בסוף הספר/פרק מצורפת הטבלה. בטבלה מופיע גרף המתאר את השכיחות היחסית של Z, והיא השטח הצבוע בגרף. הלכה למעשה, בטבלה מופיע טור עם הערכים האפשריים עבור Z

(עד ספרה אחת אחרי הנקודה) ובשורה תוספת של ספרת המאיות במידה וקיימת. בהינתן Z מחפשים בטור את Z ללא הספרה השניה אחרי הנקודה, בשורה מוצאים את המספר המתאים לספרת המאיות ומצליבים את שני הערכים. הערך המתקבל הוא α. (במידה והספרה השניה אחרי הנקודה היא אפס, מצליבים עם הערך הראשון בשורה, 00.).
כעת באמצעות כל האמור לעיל, נוכל להביט בכמה דוגמאות חישוביות.

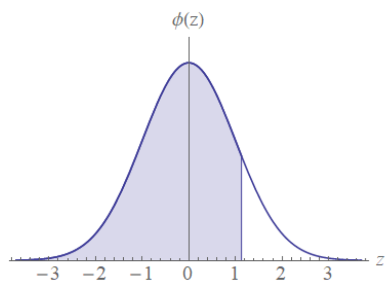
1) $\Phi(0.4) = 0.6554$

במקרה זה $Z = 0.4$ ואין ספרת מאיות, לכן מחפשים בטור את 0.4 ורואים שלידו מופיע הערך 0.6554



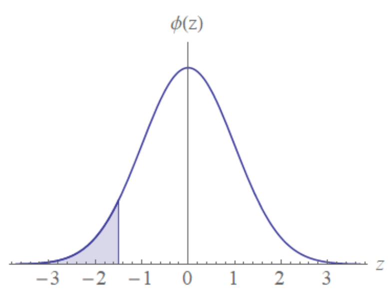
2) $\Phi(1.13) = 0.8708$

במקרה זה $Z = 1.13$ מחפשים בטור את הערך 1.1 ובשורה את 0.03 מצליבים את הערכים ומקבלים 0.8708

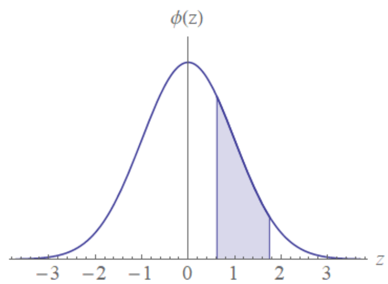


3) $\Phi(-2.5) = 1 - \Phi(2.5) = 1 - 0.9938 = 0.0062$

את $\Phi(2.5)$ מוצאים לפי השיטה לעיל. ומציבים בתכונה הראשונה הנובעת מסימטריה.



4) השכיחות היחסית בקטע $[0.62,1.75] = \Phi(1.75) - \Phi(0.62) = 0.9599 - 0.7324 = 0.2275$



הערה: עקב הסימטריות סביב 0, קל לנחש ש־ $\Phi(0) = 0.5$

<u>מציאת Z:</u>

בהינתן α, השכיחות היחסית, נתעניין בערך שכל מספר השווה לו או קטן ממנו, השכיחות לקבל אותם היא α. למעשה, אנו שואלים מה הוא Z עבורו $\Phi(Z) = \alpha$,

במקרה זה מחפשים בתוך הטבלה את α ואחרי שמוצאים, בודקים מי הם המספרים בטור ובשורה המתאימים לערך הזה, והוא Z_α. למעשה, זו הפעולה ההפוכה למציאת Z בהינתן α.

לדוגמא: $\alpha = 0.9370$ מתאים לערך $Z = 1.53$. זאת אומרת $Z_{0.9370} = 1.53$

5.4 רווחי סמך

ניתן לראות כי תחת הקירוב הנורמלי השכיחות היחסית של ממוצעים $\bar{X}_{(\mathrm{m})}$ המקיימים את התנאי

$$-Z_{1-\frac{\alpha}{2}} \leq \frac{\bar{X}_{(\mathrm{m})} - \bar{X}}{\mathrm{SD}(X)/\sqrt{m}} \leq Z_{1-\frac{\alpha}{2}}$$

היא $1 - \alpha$. תנאי זה שקול לכך ש

$$\bar{X}_{(\mathrm{m})} - \frac{\mathrm{SD}(X)}{\sqrt{m}} Z_{1-\frac{\alpha}{2}} \leq \bar{X} \leq \bar{X}_{(\mathrm{m})} + \frac{\mathrm{SD}(X)}{\sqrt{m}} Z_{1-\frac{\alpha}{2}}.$$

ולכן, אנו רואים כי הקטע $\left[\bar{X}_{(\mathrm{m})} - \frac{\mathrm{SD}(X)}{\sqrt{m}} Z_{1-\frac{\alpha}{2}}, \bar{X}_{(\mathrm{m})} + \frac{\mathrm{SD}(X)}{\sqrt{m}} Z_{1-\frac{\alpha}{2}}\right]$ מכיל את $\bar{X}$ בשכיחות יחסית של $1 - \alpha$. הקטע הנ`ל נקרא `רווח סמך ל־$\bar{X}$ ברמת סמך (או ברמת ביטחון) של $1 - \alpha$`.

5.4.1 ישום רווחי סמך

הסטטיסטיקאי או החוקר, אוחז בידו $\bar{X}_{(\mathrm{m})}$ בודד וממנו הוא מנסה להקיש על $\bar{X}$ על ידי יצירת הקטע הנ`ל. הוא איננו יודע אם הקטע מכיל את $\bar{X}$ או לא (ונדגיש, יש שתי אפשרויות בלבד: מכיל או לא מכיל). רמת הסמך טוענת שבין כל ערכי $\bar{X}_{(\mathrm{m})}$ האפשריים, בשכיחות יחסית של $1-\alpha$ הקטע יכיל את $\bar{X}$ (ובשכיחות יחסית של α לא יכיל). מה התשובה במקרה שלו, זאת איננו יודעים. למעשה רמת הביטחון אינה אומרת דבר על המקרה הנוכחי. היא טענה איכותית על השיטה בכללותה של יצירת רווחי סמך.

בעיה מעשית (אך לא תיאורתית) היא כי על מנת שניצור את רווח הסמך יש לדעת את ערכו של $\mathrm{SD}(X)$ (ושוב, ידיעת או אי ידיעת $\mathrm{SD}(X)$ אינה רלוונטית לשאלה האם הקטע מכיל את $\bar{X}$ או לא). תשובת אנשי המעשה היא כי לעיתים יש ידע מוקדם, הערכה או הנחה על ערכו של $\mathrm{SD}(X)$ ואז ניתן להשתמש בערך הנתון. כאשר אין זה כך, נחליף את $\mathrm{SD}(X)$ באומדן שלו והוא סטיית התקן במדגם עצמו, קרי $\sqrt{\frac{\sum_{i=1}^{m}(X_i - \bar{X}_{(\mathrm{m})})^2}{m}}$.

דוגמא לשימוש ברווח סמך:

בעיר כלשהי בעולם דגמו מחיר של מוצר מסויים מ־16 חנויות שונות. להלן המחירים:
95,108,97,112,99,106,105,100,99,98,104,110,107,111,103,110
ממוצע מחיר המוצר בכל המדינה אינו ידוע, אבל סטיית התקן שלו כן, והיא $SD(X)=5$.
נרצה לחשב רווח סמך לממוצע מחיר המוצר במדינה ברמת סמך (ביטחון) של 95% כלומר, $1-\alpha=0.95$. באמצעות הנוסחא הנתונה לעיל אנו יודעים שממוצע מחיר המוצר נמצא במרווח $\left[\bar{X}_{(\mathrm{m})}-\frac{\mathrm{SD}(X)}{\sqrt{m}}Z_{1-\frac{\alpha}{2}},\bar{X}_{(\mathrm{m})}+\frac{\mathrm{SD}(X)}{\sqrt{m}}Z_{1-\frac{\alpha}{2}}\right]$ ברמת סמך של 95%.
מהנתונים: $\bar{X}_{(m)}=\bar{X}_{(16)}=\frac{95+108+...+110}{16}=104$,וכן, $Z_{1-\frac{\alpha}{2}}=Z_{0.975}=1.96$ ומכאן ש: $\bar{X}\in[104-\frac{5}{4}\cdot 1.96,104+\frac{5}{4}\cdot 1.96]=[101.55,106.45]$ כלומר ממוצע מחיר המוצר במדינה נמצא ברווח הסמך $[101.55,106.45]$ ברמת סמך של 95%.

5.4.2

נחזור לדוגמא מסוף סעיף 2.2. זהו מקרה שבו האומד ל $SD(X)$ נגזר מידית מהאומד ל $\bar{X}$ והוא $\sqrt{\bar{X}_{(m)}(1-\bar{X}_{(m)})}$. נזכיר כי $\bar{X}_{(m)}$ במקרה זה הוא השכיחות היחסית של התכונה הנבדקת במדגם ו־ $\bar{X}_{(m)}(1-\bar{X}_{(m)})$ היא שונות המדגם. רווח הסמך יהיה אז:

$$\bar{X}_{(m)}-\frac{\sqrt{\bar{X}_{(m)}(1-\bar{X}_{(m)})}}{\sqrt{m}}Z_{1-\frac{\alpha}{2}}\leq\bar{X}_{(m)}\leq\bar{X}_{(m)}+\frac{\sqrt{\bar{X}_{(m)}(1-\bar{X}_{(m)})}}{\sqrt{m}}Z_{1-\frac{\alpha}{2}}$$

6 בדיקת השערות

6.1 השערת האפס וערך P

נניח כי לחוקר או לסטטיסטיקאי יש הערכה קודמת, הקרויה בקונטקסט זה השערת האפס או H_0, לגבי ערכו של $\bar{X}$. נסמנה באות היוונית μ. הוא מעוניין לתת מדד להערכה זו על סמך השוואת μ ל־$\bar{X}_{(\mathrm{m})}$. ברור כי ככל שערכים אלו יהיו רחוקים יותר, המדד אמור לתת ערכים נמוכים יותר. הוא למשל יכול להיות פונקציה של $\bar{X}_{(\mathrm{m})}-\mu$ או $\frac{\bar{X}_{(\mathrm{m})}-\mu}{\mathrm{SD}(X)/\sqrt{\mathrm{m}}}$. היתרון של האחרון הוא כי הוא חף מיחידות מדידה ולכן מנטרל השפעה של סקלה. המדד המוצע כאן הוא

$$P(\mu)=2\left(1-\Phi\left(\frac{|\bar{X}_{(\mathrm{m})}-\mu|}{\mathrm{SD}(X)/\sqrt{\mathrm{m}}}\right)\right)$$

והוא מכונה ערך P של המדגם, הוא גם מכונה סטטיסטי המבחן כי למעשה הוא הערך היחיד המעניין אותנו. נשים לב כי ערך ה־P מקבל ערכים בין אפס ל־אחת (ואכן יש לו משמעות של שכיחות יחסית). יתר על כן, ככל ש $\bar{X}-\mu$ גדול יותר, ערך P קטן יותר.
מהי משמעותו של ערך P? הוא אומר כי מכל ערכי $\bar{X}_{(\mathrm{m})}$ הבאים בחשבון, השכיחות היחסית של אילו שמרחקם מ־μ הוא כמו זה שהתקבל או אף יותר היא $P(\mu)$.
אם נחזור לנקודה של בדיקת השערות, אזי ניתן לומר כי ערך P מייצג את מידת האמון שלנו בנכונותה של השערת האפס בהתחשב בתוצאות המדגם. הרי על סמך מדגם חלקי

באוכלוסיה לעולם לא נדע האם ההשערה נכונה או לא. ככל שערך P נמוך יותר, נחשוד יותר בנכונותה של השערת האפס.

6.2 הכרעות

הדיון הנ`ל הגדיר מדד, מבוסס מדגם, המכמת את נכונותה של ההשערה כי $\mu = \bar{X}$. סטטיסטיקאי חייו קלים. מה עם מקבל ההחלטות שצריך להכריע בינרית אם השערה זו נכונה או לא. כיצד יכריע?

ראשית נאמר כיצד לא יכריע. אין אנו ממליצים כי מקבל ההחלטות יתבונן בערך $P(\mu)$ יגרד בראשו ואז יכריע. אנו מציעים אחרת: כלל הכרעה יקבע לפני ביצוע המדגם. למשל, קבע ערך קריטי, נסמנו ב־α, כך שההשערה תדחה אם $P(\mu) < \alpha$ ואחרת לא. ז`א קבענו כלל הכרעה לפני ביצוע הדגימה ואח`כ הוא הופעל הלכה למעשה.

ערכו של α, נקרא רמת המובהקות של המבחן. נכון, את α יש לקבוע מראש אך יש ערכים מקובלים. למשל, במדעי החברה $\alpha = 0.05$ מקובל. יתכן ומבט ראשון יביא לתמיהה מדוע מספר כל כך נמוך. מה רע ב־$\alpha = 0.5$? התשובה נעוצה ביישום. בדרך כלל ההשערה $\mu = \bar{X}$ מקורה בסטטוס קוו או בברירת מחדל והיינו רוצים עובדות חותכות על מנת שנחשוב ש־$\mu = \bar{X}$ כבר אינו מצב הטבע.

6.3 מבחנים חד צדדיים

עד כה השערת האפס הייתה $\mu = \bar{X}$ ודחייתה פירושו הסכמה או הכרעה כי $\mu \neq \bar{X}$. זוהי האלטרנטיבה. אך מה היה קורה אם האלטרנטיבה היתה $H_1 : \bar{X} < \mu$? נשים לב כי יש כאן הנחה נסתרת האומרת כי $\bar{X} > \mu$ אינו בא בחשבון.

כעת רק ערכים נמוכים עבור $\frac{\bar{X}_{(\mathrm{m})}-\mu}{\mathrm{SD}(X)/\sqrt{\mathrm{m}}}$ ירמזו כנגד השערת האפס. עכשיו ערך $P(\mu)$ יוגדר כ־$\Phi\left(\frac{\bar{X}_{(\mathrm{m})}-\mu}{\mathrm{SD}(X)/\sqrt{\mathrm{m}}}\right)$.

כמובן, אם האלטרנטיבה היתה $\bar{X} > \mu$ ערך $P(\mu)$ היה $1 - \Phi\left(\frac{\bar{X}_{(\mathrm{m})}-\mu}{\mathrm{SD}(X)/\sqrt{\mathrm{m}}}\right)$.

דוגמא למבחן חד־צדדי:

חקלאי מגדל תירס ולאחרונה הוא אינו מרוצה מהתבואה. הוא מעוניין להוסיף דשן על מנת לשבח את התירס. הדשן אינו זול והחקלאי מעוניין לבדוק האם משתלם לו לרכוש את הדשן. לשם כך, הוא משתמש בדשן באופן חד פעמי ולאחר מכן דוגם 25 תירסים ושוקל אותם. להלן התוצאות:

$170, 167, 174, 179, 179, 187, 179, 183, 179, 156, 163, 156, 187, 156$

$167, 156, 174, 170, 183, 179, 174, 179, 170, 159, 187$ (הנתונים הם בגרם).

ידוע כי ממוצע משקלו של קלח תירס איכותי הוא 170 גרם עם סטיית תקן של 10 גרם, ואם כך החלקאי רוצה לבדוק את ההשערה הבאה:

$$H_0 : \mu \leq 170g$$

$$H_1 : \mu > 170g$$

החקלאי בוחר רמת מובהקות של $\alpha = 0.05$. ממוצע התירסים לאחר השמת הדשן הוא: $\bar{X}_{(m)} = \bar{X}_{(25)} = 172.52$ כעת, האם משתלם לו לקנות את הדשן או לא?

כאמור, נחשב את רמת המובהקות של המבחן ועל סמך זה נקבל את ההחלטה. ערכו של P יהיה:

$$P = 1-\Phi\left(\frac{|\bar{X}_{(\mathrm{m})}-\mu|}{\mathrm{SD}(X)/\sqrt{\mathrm{m}}}\right) = 1-\Phi\left(\frac{172.52-170}{\frac{10}{5}}\right) = 1-\Phi(1.26) = 1-0.8962 = 0.1038$$

כלומר $P = 0.1038$

מה משמעות התוצאה עבורנו? $P = 0.1038 > 0.05 = \alpha$ ולכן **לא** נדחה את השערת האפס. כלומר החקלאי כנראה יעדיף שלא לרכוש את הדשן.

7 שאלות חזרה

1. לפניכם נתונים 13 אורכים שונים **במילימטרים** של גבעולי פרח הכלנית שמדד מר המיולין: $(19, 9, 32, 26, 14, 29, 12, 37, 31, 26, 19, 4, 18)$ חשבו את הממוצע, החציון וסטיית התקן של אורכי הגבעולים. ציירו הסטוגרמה של הנתונים בטווחים של חמישה מילימטרים.

2. בהמשך לנתוני השאלה הקודמת, חשבו את הממוצע, החציון וסטיית התקן של אורכי הגבעולים **בסנטימטרים**, האם ניתן להשתמש בתוצאות השאלה הקודמת?

3. מנהל בית ספר רוצה להוסיף שיעורי תגבור לתלמידי כיתות ו. התקציב של בית הספר מוגבל ומאפשר הוספת שיעורי עזר רק לאחת משלוש הכיתות שבשכבה. במבחן האחרון בחשבון התקבלו התוצאות הבאות: הממוצע בכיתה ו1 הוא 70, החציון 70 וסטיית התקן־5, בכיתה ו2 הממוצע 70החציון 65 וסטיית התקן 5. לאיזה מן הכיתות כדאי לדעתך לתת שיעורי תגבור?

4. בנק א' מציע ללקוחותיו תכנית חיסכון לשלוש שנים עם ריבית שנתית של 1% בשנה הראשונה, 5% בשנה השנייה ו־9% בשנה השלישית. בנק ב' מציע ללקוחותיו תכנית חיסכון מקבילה עם ריבית שנתית קבועה של 5% בטענה כי התכנית זהה לזו של בנק א'. האם התוכניות זהות? אם לא, מה צריכה להיות ריבית שנתית קבועה כך שהתוכניות יהיו זהות לחלוטין?

5. הוכח את הטענות הבאות או הבא דוגמא נגדית:

   (א) הממוצע תמיד גדול מהנתון המינימאלי

   (ב) הממוצע תמיד שווה לחצי מסכום הנתונים המינימאלי והמקסימאלי

   (ג) אם הממוצע שווה לחצי מסכום הנתונים המינימאלי והמקסימאלי אז הממוצע והחציון שווים

   (ד) אם הממוצע שווה לנתון המקסימאלי אז הנתון המינימאלי והמקסימאלי שווים

6. בכיתה ט' 1 ממוצע הציונים בתנ`ך הוא 73, וסטיית התקן היא $18$. לימור מכיתה ט' 1 קיבלה $85$. בכיתה ט' 2 ממוצע הציונים בתנ`ך הוא 80, וסטיית התקן היא 15. שמעון מכיתה ט' 2 קיבל 98. שמעון: הציון שלי בתנ`ך גבוה משלך, לפיכך אני הצלחתי יותר ממך ! לימור: אז מה, מעמדי בכיתה בתנ`ך גבוה ממעמדך בכיתתך, לכן בעצם, אני הצלחתי יותר ממך! מי צודק? נמקו.

7. נתונות שתי סדרות שוות אורך של מספרים: $x_1, x_2....x_n$ ו־ $y_1, y_2....y_n$. נגדיר שתי סדרות חדשות באופן הבא: $v_i = x_i y_i$ ו־ $w_i = x_i + y_i$. בכל אחד מהסעיפים הבאים מופיע מדד כלשהו (מרכז או פיזור). קבע אם מדד זה של סדרת ה־w־ים הוא סכום המדדים המתאימים של סדרת ה־x־ים וסדרת ה־y־ים, ואם מדד זה של סדרת ה־v־ים הוא מכפלת המדדים המתאימים של סדרת ה־x־ים וסדרת ה־y־ים.

אם תשובתכם היא `כן ־ תמיד`, הוכיחו. אם תשובתכם היא `לא בהכרח תמיד` תנו דוגמה נגדית.

(א) ממוצע

(ב) חציון

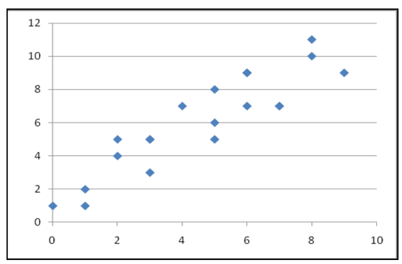
(ג) שונות

(ד) סטיית תקן

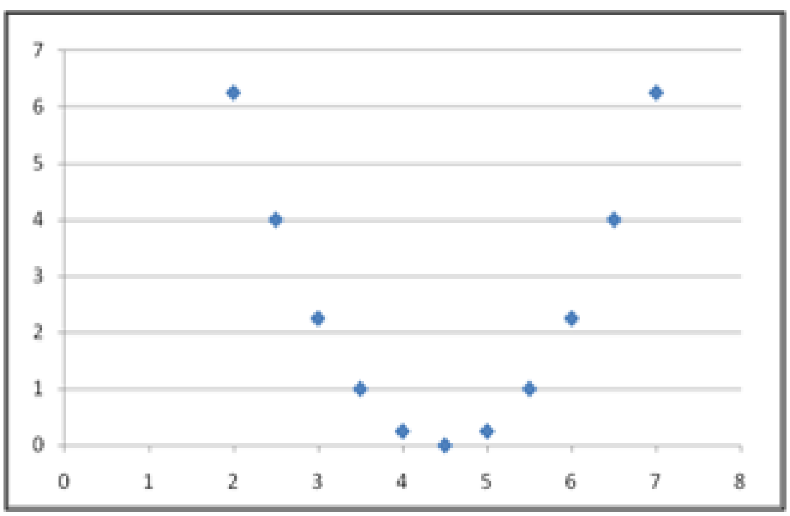
8. יהי $x_1, x_2....x_n$ אוסף נתונים כלשהו ויהיו a, b, c, d קבועים. נסמן: $y_i = ax_i + b$ ו־ $w_i = cx_i + d$ עבור $i = 1....n$. הראו כי $z_{y_i} = (y_i - \bar{y})/sd(y)$ ציון התקן של y_i, זהה ל $z_{w_i} = (w_i - \bar{w})/sd(w)$ ציון התקן של w_i. כלומר, ציוני התקן אינם משתנים עם טרנפורמציות לינאריות.

9. נתונות דיאגרמות פיזור. לפני שנותנים את הערכים ומחשבים ערך מספרי למקדם המתאם, מה לדעתכם הוא יצא? מדוע?

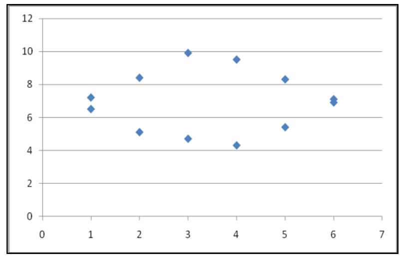
(א)



(ב)



(ג)



10. להלן זוגות של תצפיות (x_i, y_i) $i = 1...12$:

$$(6,7)(5,8)(5,6)(5,5)(4,7)(3,5)(3,3)(2,5)(2,4)(1,2)(1,1)(0,1)$$

חשבו את השונות המשותפת של x ו־ y.

11. עבור סדרת תצפיות נתון כי $\bar{x} = 3$ $\bar{y} = 3$ ו־ $\bar{xy} = 12$ מהי השונות המשותפת של x ו־ y?

12. יוסי קונה ומוכר בננות מיקרוגלים ומקפיאים. מחיר השוק של כל אחד מאלו נקבע כל יום מחדש, ולכן כל בוקר יוסי צריך ללכת לכל ספק ולשאול אותו מה מחיר המוצר שלו היום. יוסי עצלן ולכן הוא מעדיף ללכת לכמה שפחות ספקים. הוא חושב שבמקום ללכת לכל יבואן (בננות בע`מ, מיקרוגלים בע`מ ומקפיאים בע`מ) ולשאול מה המחיר של המוצר היום, מספיק לו ללכת ליבואן אחד, זאת משום שהוא חושד שהמחיר עולה ויורד בצורה מתואמת ואפשר לנבא את מחיר כל המוצרים בעזרת מחיר של רק אחד מהם. להלן נתונים של מחירי המוצרים שאסף יוסי במשך עשרה ימי עסקים:

חברה			
מקפיאים בע"מ	מיקרוגלים בע"מ	בננות בע"מ	יום עסקים
38 ש"ח	30 ש"ח	31 ש"ח	1
36 ש"ח	32 ש"ח	34 ש"ח	2
41 ש"ח	40 ש"ח	25 ש"ח	3
41 ש"ח	34 ש"ח	23 ש"ח	4
32 ש"ח	31 ש"ח	18 ש"ח	5
24 ש"ח	28 ש"ח	11 ש"ח	6
27 ש"ח	27 ש"ח	4 ש"ח	7
20 ש"ח	35 ש"ח	11 ש"ח	8
19 ש"ח	40 ש"ח	5 ש"ח	9
16 ש"ח	33 ש"ח	7 ש"ח	10

(א) ציירו גרף המתאר את המחיר של כל מוצר כפונקציה של הזמן. האם ניתן להסיק על קשר בין המחירים מתוך הדיאגרמה?

(ב) חשבו את שלושת מקדמי המתאם בין כל זוג מוצרים והחליטו איזה זוג הכי `מבטיח`, נמקו.

(ג) ציירו את דיאגרמות הפיזור של מחירי בננות ומיקרוגלים, בננות ומקפיאים ומיקרוגלים ומקפיאים, האם מקדמי המתאם שחישבתם מתארים נכון את הציורים שקיבלתם?

13. סטודנט לתואר ראשון בפסיכולוגיה רצה לבדוק את המתאם בין ציוני הפסכומטרי (טווח ציונים 200-800 ממוצע 500) לציוני מתא`ם (טווח ציונים 50-150 ממוצע 100) הוא החליט שיהיה לו יותר נוח אם ימיר את שתי הסקאלות הציונים לסקאלה אחת. לפיכך, הוא החליט להחסיר 200 נקודות מכל ציון פסיכומטרי, כך שטווח הציונים החדש יהיה 0-600 וממוצע 300, ולהחסיר 50 נקודות מכל ציון מתא`ם ואז להכפילו פי 6 כך שטווח הציונים החדש של המתא`ם יהיה אף הוא 0-600 וממוצע 300. כיצד שינוי הנתונים ישפיע על מקדם המתאם?

14. נתונה טבלת ציונים של 21 סטודנטים בשני מקצועות – היסטוריה וכימיה:

סטודנט	כימיה	היסטוריה
1	85	65
2	74	50
3	76	55
4	90	65
5	85	55
6	87	70
7	94	65
8	98	70
9	81	55
10	91	70
11	76	50
12	74	55

(א) ציירו את הנתונים בדיאגרמת פיזור כאשר ציר ה-X הוא הציון בהיסטוריה וציר ה-Y הוא הציון בכימיה, וחשבו את ישר הרגרסיה. ציירו את ישר הרגרסיה של Y על X.

(ב) אהוד קיבל 70 בהיסטוריה, אך החסיר את הבחינה בכימיה בשל מחלה. איזה ציון אופייני לסטודנט כזה לפי הנתונים ?

(ג) א. המורה להיסטוריה החליט לבטל שאלה אחת בבחינה לכן כל תלמיד שיפר את ציונו ב 15% - כיצד שינוי זה ישפיע על מקדמי הרגרסיה?
ב. המורה לכימיה גילה כי התלמידים העתיקו לכן הוא הוריד להם 10% מהציון - כיצד שינוי זה ישפיע על מקדמי הרגרסיה?
ג. הוכיחו מסקנה כללית לגבי אופן ההשפעה של טרנספורמציה לינארית של הנתונים (הכפלה בקבוע והוספה של קבוע) על מקדמי הרגרסיה. הפרידו למקרים שהטרנספורמציה היא ב-X-ים וב-Y-ים

(ד) חשבו את הקורלציה ואת R בריבוע. בעזרת הקורלציה מצאו את קו הרגרסיה של ציוני התקן של Y על ציוני התקן של X ללא חישובים נוספים.

(ה) חשבו את ציוני התקן של האדם הראשון בקובץ. מה הערך החזוי של אדם זה במונחי ציוני התקן (דהיינו, מה הערך החזוי של ציון התקן בכימיה על פי ציון התקן בהיסטוריה אצל אדם זה). איזו תפועה באה כאן לידי ביטוי?

15. רוצים להתאים לנתונים קו רגרסיה מהצורה $y = ax$ (ללא חותך, הקו יוצא מראשית הצירים)

(א) חשבו את פתרון הריבועים הפחותים לבעיה, כלומר, מצאו ערך a המקיים: $\min \sum_{i=1}(y_i - ax_i)^2$

(ב) האם הקו הנ`ל עובר בנקודת הממוצעים?

16. עבור הדוגמא עם החקלאי וקלחי התירס מ־6.3 מצאו רווח סמך ל μ עבור:

(א) $\alpha = 0.05$

(ב) $\alpha = 0.01$

17. מה צריך להיות גודלו של מדגם מסויים, m , בהינתן שרווח הסמך שלו ל μ הינו ברמת סמך של 95% ורוחב רווח הסמך, ε מקיים $\varepsilon \leq 1.92$?

18. משרד הבריאות פרסם כי עבור נשים בוגרות מתחת לגיל 51, הקצבה היומית המומלצת של ברזל היא $18mg$. חוקר דגם את רמת הברזל (X) אצל 45 נשים מתחת לגיל 51 במשך יממה. תוצאות המדגם היו: $\bar{X}_{(45)} = 14.68mg$. ונתון: $SD(X) = 4.2mg$.
ברמת מובהקות של $\alpha = 0.01$ בדקו את המבחנים הבאים:

(א) $H_0 : \mu = 18mg$ למול $H_1 : \mu \neq 18mg$

(ב) $H_0 : \mu = 18mg$ למול $H_1 : \mu < 18mg$

חלק II

מבוא לתורת הקבוצות ולפונקצית הסתברות

8 קבוצות ופעולות על קבוצות

הסטטיסטיקה התאורית שבה עסקנו עד עתה קובעת כלים יעילים לניתוח מאפיינים של קבוצות נתונים. למשל ראינו שהממוצע החשבוני של סדרת נתונים מביא למינימום את פונקציית המרחק של סכום ריבועי הסטיות. זו עובדה מתמטית טהורה שאינה קשורה בהכרח לטבע העולם ולכן כשלעצמה היא לא עוזרת לנו להסיק כל מסקנה מדויקת. כדי להסיק מסקנות, נצטרך להשתמש ב**מודל**. נעסוק במודל הנפוץ של **תורת ההסתברות**. נציין רק כי המסקנות נכונות תמיד במודל אבל לא בהכרח במציאות.

לשם כך נציג תחילה מבוא שיכיל מושגים כלליים ויסודיים מתורת הקבוצות (נציג את המונחים בהקשר ובשפה של סטטיסטיקה, אולם למעשה מה שנראה בפרק 2 אלה מושגים כלליים בתורת הקבוצות), ולאחר מכן נציג את המודל של תורת ההסתברות.

8.1 מונחים יסודיים

1. מבצעים ניסוי כלשהו.]למשל הטלת קוביה.[כל אחת מהתוצאות האפשריות נקראת `**מאורע פשוט**` ומסומנת ב–ω_i.]תוצאה אפשרית בדוגמה שלנו היא 2 או 4.[

2. אוסף כל התוצאות האפשריות נקרא `**מרחב המדגם**` ומסומן Ω (אומגה).]בדוגמה זו מרחב המדגם הוא $\{1,2,3,4,5,6\}$[. כלומר, אם יש n אפשרויות אז $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$.]בדוגמה זו יש 6 אפשרויות.[

3. אוסף כלשהו של תוצאות אפשריות נקרא `**מאורע**`.]למשל $\{1,3\}$ ו–$\{2,3,4,5\}$ הם מאורעות.[נשים לב ש–Ω כולה היא סוג של מאורע, כי היא אוסף כלשהו של תוצאות אפשריות.

4. נסמן את הקבוצה הריקה של המאורעות ב–$\emptyset$. קבוצה זו היא ה`אפס` של המאורעות.

שייכות: נשתמש בסימן `$\in$` כדי לקבוע שמאורע פשוט שייך למאורע. כך למשל המאורע הפשוט `3` שייך למאורע $\{1,3\}$, ולכן נסמן $3 \in \{1,3\}$. לעומת זאת המאורע הפשוט `7` אינו שייך למאורע $\{1,3\}$ ולכן נסמן $7 \notin \{1,3\}$.

הכלה: נאמר שמאורע A מוכל במאורע B, אם לכל $a \in A$ מתקיים גם $a \in B$. כדי לציין שמאורע A מוכל במאורע B נסמן $A \subseteq B$. נשים לב שהמאורע $\emptyset$ מוכל בכל מאורע, וכן שכל מאורע מוכל במאורע Ω.

שוויון: נאמר שמאורעות A, B שווים אם מתקיים $A \subseteq B$ וגם $B \subseteq A$.

מאורע משלים: נאמר שמאורע B הוא המשלים של מאורע A, אם הוא מכיל את כל המאורעות הפשוטים השייכים ל–Ω אך לא ב–A. כך למשל בניסוי של הטלת קוביה, המאורע $B = \{1,2,6\}$ הוא המשלים של המאורע $A = \{3,4,5\}$. נסמן ב–$\overline{A}$ את המאורע המשלים של A.
נשים לב שמתקיים:

1. $\overline{\Omega} = \emptyset$, $\Omega = \overline{\emptyset}$ לכל מרחב מדגם.
2. $\overline{\overline{A}} = A$ לכל מאורע.
הערה: משמעותו של שוויון זה היא שמתקיימת סימטריה. כלומר, אם B הוא מאורע משלים של A, אז A הוא מאורע משלים של B.

איחוד מאורעות: איחוד המאורעות A, B הוא מאורע שמכיל את כל המאורעות הפשוטים ששייכים ל־A או ששייכים ל־B. (`או` במשמעותו המתמטית. כלומר, כולל המאורעות הפשוטים ששייכים לשניהם). נסמן את איחוד המאורעות A, B ב־$A \cup B$. למשל $\{1,2\} \cup \{2,4,5\} = \{1,2,4,5\}$.
נשים לב שמתקיים:

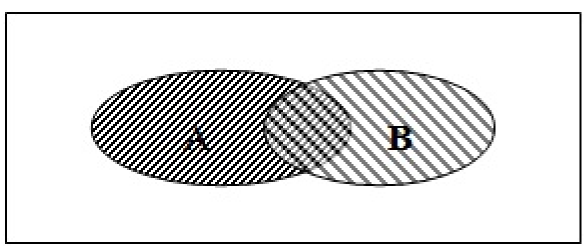
1. $A \cup \overline{A} = \Omega$ לכל מאורע A.
2. $A \cup \emptyset = A$ לכל מאורע A.

חיתוך מאורעות: חיתוך המאורעות A, B הוא מאורע שמכיל את המאורעות הפשוטים ששייכים ל־A וגם ל־B. באופן פורמלי, $x \in A \cap B$ אם $x \in A$ וגם $x \in B$. נסמן את חיתוך המאורעות A, B ב־$A \cap B$. למשל $\{1,2\} \cap \{2,4,5\} = \{2\}$.
נשים לב שמתקיים:

1. $A \cap \overline{A} = \emptyset$ לכל מאורע A.
2. $A \cap \emptyset = \emptyset$ לכל מאורע A.
3. $A \cap \Omega = A$ לכל מאורע A.

מאורעות זרים: המאורעות A, B נקראים זרים אם $A \cap B = \emptyset$.

דיאגרמת־ון: דיאגרמות מסוג זה שנתאר מיד, הן כלי שימושי אך לא־פורמלי להבנת היחסים של שייכות, הכלה, איחוד, חיתוך והשלמה שבין מאורעות שונים. דיאגרמת־ון כללית של מאורעות המסומנים A, B נראית כך:



כאשר המלבן כולו מייצג את מרחב המדגם Ω, ושני העיגולים מייצגים שני מאורעות A, B. השטח החופף לשתי האליפסות מייצג את החיתוך $A \cap B$. השטח של שתי האליפסות, כאשר את השטח החופף מחשבים פעם אחת, מייצג את האיחוד $A \cup B$. השטח הכולל של המלבן פחות השטח של שתי האליפסות, מייצג את המשלים $\overline{A \cup B}$. וכן באופן דומה ניתן לסמן מאורעות נוספים בדיאגרמה וליצור יחסים אחרים.

דוגמה: נתייחס בדוגמה זו לדיאגרמה שהוצגה לעיל. נגדיר את מרחב המדגם Ω להיות כל הסטודנטים והסטודנטיות משנה א'. נגדיר את המאורע A להיות הבנים, כך שהמאורע $\overline{A}$ הוא הבנות. נגדיר את המאורע B להיות הסטודנטים והסטודנטיות בעלי העיניים הכחולות, והמאורע $\overline{B}$ להיות כל השאר. לפי הגדרות אלה המאורע $A \cap B$ הוא כל הבנים בעלי העיניים הכחולות. כמו־כן המאורע $A \cup B$ הוא כל הבנים, בתוספת הבנות בעלות העיניים הכחולות. או באופן שקול: קבוצות הסטודנטים והסטודנטיות בעלי העיניים הכחולות, בתוספת הבנים בעלי עיניים שאינן כחולות. נשים לב שמתקיים $\overline{A \cup B} = \overline{A} \cap \overline{B}$. כלומר, כל הבנות בעלות עיניים שאינן כחולות.

8.2 כללי דה־מורגן

טענה 1: לכל שתי קבוצות A, B מתקיים:

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

הוכחה: נוכיח את השוויון באמצעות הכלה דו־כיוונית.

- כיוון ראשון: יהי w מאורע פשוט כלשהו המקיים $w \in \overline{A \cup B}$. נסיק:

$$\begin{gathered} w \in \overline{A \cup B} \\ \Downarrow \\ w \notin A \cup B \\ \Downarrow \\ w \notin A \,\text{and}\, w \notin B \\ \Downarrow \\ w \in \overline{A} \,\text{and}\, w \in \overline{B} \\ \Downarrow \\ w \in \overline{A} \cap \overline{B} \end{gathered}$$

ולכן $\overline{A \cup B} \subseteq \overline{A} \cap \overline{B}$.

- כיוון שני: יהי w מאורע פשוט כלשהו המקיים $w \in \overline{A} \cap \overline{B}$. נסיק:

$$\begin{gathered} w \in \overline{A} \cap \overline{B} \\ \Downarrow \\ w \in \overline{A} \,\text{and}\, w \in \overline{B} \\ \Downarrow \\ w \notin A \,\text{and}\, w \notin B \\ \Downarrow \\ w \notin A \cup B \\ \Downarrow \\ w \in \overline{A \cup B} \end{gathered}$$

ולכן $\overline{A} \cap \overline{B} \subseteq \overline{A \cup B}$.

- נסיק משני הכיוונים שלפי הגדרת השוויון מתקיים $\overline{A} \cap \overline{B} = \overline{A \cup B}$. ■

הערה: נשים לב שבשני הכיוונים ביצענו את אותם היסקים, רק בכיוונים לוגיים הפוך. כלומר כל צעד בהוכחה מהווה שקילות ולא רק גרירה בכיוון אחד, כך שיכולנו לרשום בקיצור פעם אחת את אותם שלבים לוגיים עם הסימון $\Leftrightarrow$.

טענה 2: לכל שתי קבוצות A, B מתקיים:

$$\overline{A} \cup \overline{B} = \overline{A \cap B}$$

הוכחה: נשתמש בתוצאה שהראינו בטענה הקודמת, ונסיק:

$$\begin{gathered}
\overline{A \cup B} = \overline{A} \cap \overline{B} \\
\Downarrow \\
\overline{\overline{A \cup B}} = \overline{\overline{A} \cap \overline{B}} \\
\Downarrow \\
A \cup B = \overline{\overline{A} \cap \overline{B}} \\
\Downarrow \\
\overline{A} \cup \overline{B} = \overline{\overline{\overline{A}} \cap \overline{\overline{B}}} \\
\Downarrow \\
\overline{A} \cup \overline{B} = \overline{A \cap B}
\end{gathered}$$

$\blacksquare$

הגרירה הלפני האחרונה נובעת מכך שהטענה הנכונה לכל שתי קבוצות A, B, נכונה גם עבור הקבוצות $\overline{A}, \overline{B}$. כלומר ביצענו הצבה של $\overline{A}, \overline{B}$ בשוויון הלפני אחרון שקיבלנו על A, B.

8.3 שכיחות יחסית

נניח כי נתון מרחב המדגם של הטלת קוביה: $\Omega = \{1, 2, 3, 4, 5, 6\}$. ביצענו את הניסוי 100 פעמים והתקבלו התוצאות הבאות:

תוצאות	1	2	3	4	5	6
שכיחות	15	20	25	15	10	15
שכיחות יחסית	0.15	0.2	0.25	0.15	0.1	0.15

נגדיר את f להיות פונקציה שמחזירה את השכיחות היחסית. כלומר $f(A)$ היא השכיחות היחסית של מאורע A כלשהו. ראשית נשים לב שמתקיים:

$$f\left(\emptyset\right) = 0$$

$$f\left(\Omega\right) = 1$$

נבחן למשל את המאורעות הבאים:

$$A = \{1, 2, 3\} \qquad f\left(A\right) = 0.6$$

$$B = \{2, 4, 6\} \qquad f\left(B\right) = 0.5$$

$$A \cup B = \{1, 2, 3, 4, 6\} \qquad f\left(A \cup B\right) = 0.9$$

חשוב לשים לב כי $f(A\cup B)\neq f(A)+f(B)$. הסיבה לכך היא שמאורע פשוט ששייך גם ל-A וגם ל-B נספר פעם אחת בלבד כאשר מחשבים את השכיחות היחסית של $A\cup B$. בעוד שבביטוי $f(A)-f(B)$ הוא נכנס פעמיים. מנימוק זה נסיק שמתקיימת הנוסחה:

$$\boxed{f(A\cup B)=f(A)+f(B)-f(A\cap B)}$$

- במקרה שבו המאורעות A,B זרים, מתקיים השוויון:

$$f(A\cup B)=f(A)+f(B)$$

כי $f(A\cap B)=f(\emptyset)=0$ במקרה זה האיחוד בין A ל-B יקרא `איחוד זר`.

- מכאן נוכל להסיק שמתקיים:

$$f(\overline{A})=1-f(A)$$

כי $A,\overline{A}$ הם מאורעות זרים, ולכן:

$$f(A)+f(\overline{A})=f(A\cup\overline{A})=f(\Omega)=1$$

- עבור שלושה מאורעות מתקיים:

$$f(A\cup B\cup C)=$$

$$f(A)+f(B)+f(C)-f(A\cap B)-f(A\cap C)-f(B\cap C)+f(A\cap B\cap C)$$

-

$$f(A)=f(A\cap B)+f(A\cap\overline{B})$$

הסיבה לכך היא שמתקיימים השוויונים:

$$A=(A\cap B)\cup(A\cap\overline{B})$$

$$(A\cap B)\cap(A\cap\overline{B})=\emptyset$$

- נשים לב שהמאורע $A\cap\overline{B}$ הוא `$A$ פחות $B$`. כלומר מכיל את המאורעות הפשוטים של A, למעט אלו מהם ששייכים גם ל-B, ולכן:

$$f(A\cap\overline{B})=f(A)-f(A\cap B)$$

8.4 חלוקה

הגדרה: נניח שנתון מרחב מדגם Ω כלשהו. נאמר שקבוצה של מאורעות $\{B_k\}_{k=1}^{n}$ היא חלוקה של Ω, אם מתקיימים שני תנאים הבאים:

1. המאורעות $\{B_k\}_{k=1}^{n}$ **זרים בזוגות**. כלומר, לכל $i \neq j$ עבור $i, j = 1, ..., n$ מתקיים $B_i \cap B_j = \phi$.[5]

2. המאורעות $\{B_k\}_{k=1}^{n}$ **מכסים** את Ω. כלומר, מתקיים כי $\bigcup\limits_{i=1}^{n} B_i = B_1 \cup B_2 \cup ... \cup B_n = \Omega$.

נשים לב שלכל מאורע A מתקיים כי הזוג $A, \overline{A}$ הוא חלוקה.

- נכליל את השוויון $f\left(A\right) = f\left(A \cap B\right) + f\left(A \cap \overline{B}\right)$ שראינו לעיל, לחלוקה $\{B_k\}_{k=1}^{n}$.

9 פונקציית הסתברות

נבנה מודל שמעניק משמעות פורמלית למידת הוודאות להתרחשותם של מאורעות. נציין מייד כי ודאות זו יכולה להיות סוביקטיבית אבל נקבע עבורה כמה כללים תוחמים. לשם כך נגדיר כי המאורע Ω הוא ודאי ומקבל את הערך המקסימלי 1, והמאורע $\emptyset$ יקבל את הערך המינימלי 0. כל שאר המאורעות יקבלו ערכי ביניים. נגדיר את הפונקציה P (probability) שתחזיר את ערך הוודאות של כל מאורע A. כלומר:

$$\mathrm{P}\left(\Omega\right) = 1$$

$$\mathrm{P}\left(\phi\right) = 0$$

$$0 \leq \mathrm{P}\left(A\right) \leq 1$$

דוגמה: נדון בהטלת קוביה. מניחים שסדר התוצאות אינו משנה וכן תוצאה שחוזרת על עצמה היא אותה תוצאה. נבדוק מהו מספר המאורעות האפשריים:
אם $\Omega = \{1\}$ יש שני מאורעות אפשריים: $\emptyset, \Omega$.
אם $\Omega = \{1, 2\}$ יש ארבעה מאורעות אפשריים: $\emptyset, \{1\}, \{2\}, \Omega$.
אם $\Omega = \{1, 2, 3\}$ יש שמונה מאורעות אפשריים: $\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \Omega$.
וכן הלאה...

טענה: בניסוי בעל n מאורעות פשוטים, מספר המאורעות האפשריים הוא 2^n. הדבר נובע מכך שעבור כל מאורע פשוט קיימות שתי אפשרויות: שייך למאורע או לא שייך לו.

אם־כך במקרה של הטלת קוביה פעמיים יש לנו מרחב מדגם בן $6^2 = 36$ מאורעות, ונרצה להעניק לכל מאורע מספר שיעניק ביטוי פורמלי למידת הוודאות שמאורע זה יתרחש. כפי שהגדרנו כבר לעיל, לא ייתכן שלא יקרה כלום ולכן $P\left(\emptyset\right) = 0$. כמו־כן בוודאות מאורע כלשהו מתוך מרחב המדגם יקרה ולכן $\mathrm{P}\left(\Omega\right) = 1$. כל שאר המאורעות הם במידה של ודאות שנמצאת בין 0 ל־1. נדרוש שהערכים שניתן למאורעות יהיו `הגיוניים`, במובן זה למשל אם $A \subseteq B$ אז $\mathrm{P}\left(A\right) \leq \mathrm{P}\left(B\right)$.

[5] נשים לב כי אם המאורעות זרים בזוגות אז הם זרים. כלומר מתקיים גם $B_1 \cap B_2 \cap ... \cap B_n = \emptyset$. ההיפך לא נכון.

פונקציית הסתברות

נאמר שפונקציה P כלשהי המתאימה מספר ממשי לכל מאורע במרחב המדגם, נקראת `פונקציית הסתברות` אם היא מקיימת את התנאים הבאים:

1. $\mathrm{P}(\Omega)=1$
2. לכל $A\subseteq\Omega$ מתקיים $0\leq \mathrm{P}(A)$.
3. אם A,B מאורעות זרים, אז:

$$\mathrm{P}(A\cup B)=\mathrm{P}(A)+\mathrm{P}(B)$$

שלושת התנאים האלה נקראים `אקסיומות פונקציית ההסתברות`. האקסיומות הללו בלתי תלויות. כלומר, כל שתיים מהן לא גוררות את השלישית. או באופן שקול: עבור כל שתי אקסיומות, קיימת פונקציה שאינה פונקציית הסתברות, כי היא מקיימת את השתיים הללו ולא מקיימת את השלישית. בינתיים אנו לא יודעים האם קיימת פונקציה שאכן מקיימת את שלושת התנאים הללו. מיד נראה קיום של פונקציה כזאת באמצעות דוגמה.

עוצמה של מאורע: נגדיר עוצמה של מאורע A להיות מספר המאורעות הפשוטים השייכים ל-A, ונסמן אותה ב-$|A|$. לדוגמה, עבור $A=\{1,2\}$, $|A|=2$, ועבור $\Omega=\{1,2,...,6\}$, $|\Omega|=6$.

דוגמא 1

הפונקציה $\mathbf{P}(A)=\frac{|A|}{|\Omega|}$

נבחן כעת את הפונקציה $\mathrm{P}(A)=\frac{|A|}{|\Omega|}$ ונראה שהיא פונקציית הסתברות:

1. $\mathrm{P}(\Omega)=\frac{|\Omega|}{|\Omega|}=1$
2. $0\leq\frac{|A|}{|\Omega|}\leq 1$ כי $|A|,|\Omega|>0$, וכן $|A|\leq|\Omega|$.
3. נניח כי A,B מאורעות זרים, אז אכן מתקיים:

$$\mathrm{P}(A\cup B)=\frac{|A\cup B|}{|\Omega|}=\frac{|A|+|B|}{|\Omega|}=\frac{|A|}{|\Omega|}+\frac{|B|}{|\Omega|}=\mathrm{P}(A)+\mathrm{P}(B)$$

דוגמא 2

פונקציה זו אינה הפונקציה היחידה שמקיימת את אקסיומות פונקציית ההסתברות. נגדיר את מרחב המדגם $\Omega=\{\omega_1,\omega_2,...,\omega_n\}$ ונתבונן בפונקציה החלופית הבאה:

$$\mathrm{P}(A)=\begin{cases}1 & \omega_1\in A\\ 0 & \omega_1\notin A\end{cases}$$

נראה שמתקיימים שלושת אקסיומות פונקציית ההסתברות:

1. $\omega_1 \in \Omega$ ולכן $\mathrm{P}(\Omega) = 1$.

(א) כל ערכי הפונקציה האפשריים הם $0,1$ ולכן $0 \le \mathrm{P}(A)$.

(ב) נניח כי A, B מאורעות זרים. נבדוק שלוש אפשרויות:

i. אם $\omega_1 \notin A \cup B$ (לא שייך ל-A ולא ל-B) אז ההסתברויות כולן מקיימות:

$$\mathrm{P}(A) = \mathrm{P}(B) = \mathrm{P}(A \cup B) = 0$$

ולכן מתקיימת האקסיומה.

ii. אם $\omega_2 \in A \cap \overline{B}$ (שייך ל-A ולא ל-B) אז $\mathrm{P}(A) = 1$, $\mathrm{P}(B) = 0$, $\mathrm{P}(A \cup B) = 1$ ולכן מתקיימת האקסיומה.

iii. אם $\omega_1 \in \overline{A} \cap B$ (שייך ל-B ולא ל-A) מדובר במקרה סימטרי למקרה (b).

דוגמא 3

ניתן לראות כי פונקצית השכיחות היחסית שהודגמה בפרק הקודם מקיימת את שלוש האקסיומות הללו. כמובן, זה נכון לכל פונקצית שכיחות יחסית אפשרית ולא רק לזו שנתקבלה במדגם.

מדוגמאות אלו נוכל להסיק שגם לאחר ניסוח אקסיומות פונקציית ההסתברות, נותר שיקול דעת בידי מתכנן המודל.

תכונות פונקציית ההסתברות

1. $\mathrm{P}(\emptyset) = 0$.

<u>הוכחה</u>: נשים לב שמתקיים לכל A כי $A = A \cup \emptyset$, ולכן:

$$\mathrm{P}(A) = \mathrm{P}(A \cup \emptyset) = \mathrm{P}(A) + \mathrm{P}(\emptyset)$$
$$\Downarrow$$
$$\mathrm{P}(\emptyset) = 0$$

2. אם A, B, C מאורעות זרים בזוגות, ז`א $A \cap B = A \cap B = B \cap C = \emptyset$, אז:

$$\mathrm{P}(A \cup B \cup C) = \mathrm{P}(A) + \mathrm{P}(B) + \mathrm{P}(C)$$

<u>הוכחה</u>:

$$\mathrm{P}(A \cup B \cup C) = \mathrm{P}((A \cup B) \cup C)$$

$$= \mathrm{P}(A \cup B) + \mathrm{P}(C) = \mathrm{P}(A) + \mathrm{P}(B) + \mathrm{P}(C)$$

3. לכל זוג מאורעות A, B מתקיים:

$$\mathrm{P}(A \cap B) = \mathrm{P}(A) - \mathrm{P}(A \cap \overline{B})$$

<u>הוכחה</u>: נשים לב שמתקיים $A = (A \cap B) \cup (A \cap \overline{B})$ וזה איחוד זר. מכאן:

$$\mathrm{P}(A) = \mathrm{P}((A \cap B) \cup (A \cap \overline{B})) = \mathrm{P}(A \cap B) + \mathrm{P}(A \cap \overline{B})$$

4. אם $\{B_k\}_{k=1}^{m}$ היא חלוקה של Ω, אז:

$$\mathrm{P}(A)=\sum_{i=1}^{m}\mathrm{P}(A\cap B_i)$$

הוכחה: כפי שראינו לעיל לכל A מתקיים:

$$A=(A\cap B_1)\cup(A\cap B_2)\cup...\cup(A\cap B_m)$$

כאשר זה איחוד זר בזוגות. ולכן:

$$\mathrm{P}(A)=\mathrm{P}((A\cap B_1)\cup(A\cap B_2)\cup...\cup(A\cap B_m))$$

$$=\mathrm{P}(A\cap B_1)+\mathrm{P}(A\cap B_2)+...+\mathrm{P}(A\cap B_m)=\sum_{i=1}^{m}\mathrm{P}(A\cap B_i)$$

5. לכל A,B מתקיים:

$$\mathrm{P}(A\cup B)=\mathrm{P}(A)+\mathrm{P}(B)-\mathrm{P}(A\cap B)$$

הוכחה: נשים לב שמתקיים:

$$A\cup B=\left(A\cap\overline{B}\right)\cup\left(\overline{A}\cap B\right)\cup(A\cap B)$$

וזה איחוד זר בזוגות. ולכן:

$$\mathrm{P}(A\cup B)=\mathrm{P}\left(\left(A\cap\overline{B}\right)\cup\left(\overline{A}\cap B\right)\cup(A\cap B)\right)$$

$$=\mathrm{P}\left(A\cap\overline{B}\right)+\mathrm{P}\left(\overline{A}\cap B\right)+\mathrm{P}(A\cap B)$$

$$=[\mathrm{P}(A)-\mathrm{P}(A\cap B)]+[\mathrm{P}(B)-\mathrm{P}(A\cap B)]+\mathrm{P}(A\cap B)$$

$$=\mathrm{P}(A)+\mathrm{P}(B)-\mathrm{P}(A\cap B)$$

[נשים לב שהשוויון השלישי נובע מתכונה 3.]

פונקציית הסתברות כללית: כדי לבנות פונקציה כלשהי שמקיימת את אקסיומות פונקציית ההסתברות על התחום שנסמן $\Omega=\{\omega_1,\omega_2,...,\omega_n\}$, מספיק לבנות אותה כך שתקיים את שני התנאים הבאים:

$$0\leq\mathrm{P}(\omega_i)\ ,\ 1\leq i\leq n$$

$$\sum_{i=1}^{n}\mathrm{P}(\omega_i)=1$$

לא נוכיח כאן, אולם שני תנאים אלה מספיקים בכדי להפוך בהכרח את הפונקציה לפונקציית הסתברות, המקיימת עבור מאורע כלשהו A שההסתברות היא:

$$\mathrm{P}(A)=\sum_{\omega_i\in A}\mathrm{P}(\omega_i)$$

פונקצית הסתברות אחידה: נאמר שפונקצית הסתברות היא אחידה, אם לכל אחד מהמאורעות הפשוטים שבו הסתברות שווה. כמובן, הסתברות זאת תהיה $\frac{1}{|\Omega|}$.

דוגמאות:

- **הטלת קוביה**: מרחב המדגם הוא $\Omega = \{1,2,3,4,5,6\}$ וההסתברויות הן:

$$\mathrm{P}(1) = \mathrm{P}(2) = \mathrm{P}(3) = \mathrm{P}(4) = \mathrm{P}(5) = \mathrm{P}(6) = \frac{1}{6}$$

נבחר למשל את המאורע $A = \{2,1,4\}$. ההסתברות היא $\frac{1}{2}$ כי מאורע זה כולל מחצית מהאפשרויות במרחב המדגם. ניתן לחשב גם לפי תכונות פונקציית ההסתברות:

$$\mathrm{P}(A) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

- **הטלת מטבע**: מרחב המדגם הוא $\Omega = \{H,T\}$ וההסתברויות הן $\mathrm{P}(H) = \mathrm{P}(T) = \frac{1}{2}$.

נדון למשל במקרה בו שני שחקנים מחליטים לזרוק את המטבע פעמיים, ולהגדיר:
– אם יש אפס פעמים H אז שחקן א' מנצח.
– אם יש פעם אחת H שחקן ב' מנצח.
– אם יש פעמיים H תיקו.

נשים לב לתוצאות האפשריות והסתברותן:

$\{T,T\}$	$\{T,H\}$	(H,T)	$\{H,H\}$
0.25	0.25	0.25	0.25

מכאן שהאפשרות של תיקו מתקבלת בשני אירועים, ולכן ההסתברות לתיקו היא 0.5, כפול מההסתברויות ששחקן א' ינצח וששחקן ב' ינצח. נשים לב כי במודל שהגדרנו מרחב המדגם $\Omega = \{\text{"player A wins"}, \text{"player b wins"}, \text{"tie"}\}$ אינו מוליך לפונקצית הסתברות אחידה.

10 שאלות חזרה

1. מאוכלוסיית הסטודנטים מהחוג לסטטיסטיקה נבחר באקראי סטודנט יחיד נגדיר את המאורעות:

   - A - הסטודנט קורא ישראל היום.
   - B - הסטודנט קורא ידיעות אחרונות.
   - C - הסטודנט קורא הארץ.

   (א) השתמשו בפעולות `איחוד`, `חיתוך` ו-`משלים` בלבד, בכדי לבטא בסימונים של קבוצות את המאורעות הבאים.

   i. הסטודנט קורא ישראל היום.
   ii. הסטודנט קורא ישראל היום וידיעות.
   iii. הסטודנט קורא ישראל היום, ידיעות והארץ.

iv. הסטודנט קורא ישראל היום וידיעות, אך אינו קורא הארץ.
v. הסטודנט אינו קורא אף עיתון.
vi. הסטודנט קורא בדיוק שני עיתונים.
vii. הסטודנט קורא לפחות שני עיתונים.
viii. הסטודנט קורא בדיוק עיתון אחד.
ix. הסטודנט קורא לפחות עיתון אחד.
x. הסטודנט אינו `קורא רק הארץ`.

(ב) נסחו במלים את המאורעות הבאים:

i. $B \cup C$
ii. $\overline{A \cup B}$
iii. $\left(A \cap B \cap \overline{C}\right) \cup \left(A \cap \overline{B} \cap C\right) \cup \left(A \cap B \cap C\right)$

(ג) ענו על השאלות הבאות לגבי המאורעות שהוגדרו בסעיף א':

i. האם $i \subseteq ii$? להפך? לא זה ולא זה?
ii. אם $iii \subseteq iv$? להפך? לא זה ולא זה?
iii. האם $viii \subseteq ix$? להפך? לא זה ולא זה?

2. נתון מרחב המדגם הבא: $\Omega = \{e, f, g, h\}$.

(א) רשמו את כל הקבוצות החלקיות של Ω.

(ב) נגדיר את שלושת הקבוצות הבאות: $A = \{e, f, g\}$, $B = \{e, f\}$, $C = \{g, h\}$.
הציגו את המאורעות הבאים באמצעות המאורעות הפשוטים של מרחב המדגם:

i. $A \cup B$
ii. $A \cap C$
iii. $\overline{A} \cup B$
iv. $\left(A \cup \overline{B}\right) \cap C$

3. הוכיחו באינדוקציה את כללי דה מורגן עבור כל אוסף של מאורעות $A_1, A_2, ..., A_n$, כאשר $n \geq 2$ הוא מספר שלם כלשהו:

(א) $\overline{\cup_{i=1}^{n} A_i} = \cap_{i=1}^{n} \overline{A_i}$
(ב) $\overline{\cap_{i=1}^{n} A_i} = \cup_{i=1}^{n} \overline{A_i}$

4. נתון מרחב המדגם הבא: $\Omega = \{a, b, c, d, e, f, g, h\}$ קבעו עבור כל אחת מקבוצות המאורעות הבאות האם מדובר בחלוקה של מרחב המדגם. נמקו.

(א) $A = \{a, b, c\}$ $B = \{c, d, e\}$ $C = \{e, f, g, h\}$
(ב) $A = \{a, b, c\}$ $B = \{d, e\}$ $C = \{f, g\}$
(ג) $A = \{a, b, c\}$ $B = \{d, e\}$ $C = \{f, g, h\}$
(ד) $A = \{a, c, h\}$ $B = \{d, e\}$ $C = \{b, f, g\}$

5. נתון מרחב המדגם הבא: $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, עבורו נתונות ההסתברויות הבאות:

$$\mathrm{P}\left(\{\omega_1\}\right) = 0.1, \mathrm{P}\left(\{\omega_2\}\right) = 0.1, \mathrm{P}\left(\{\omega_3\}\right) = 0.2$$
$$\mathrm{P}\left(\{\omega_4\}\right) = 0.25\mathrm{P}\left(\{\omega_5\}\right) = 0.3, \mathrm{P}\left(\{\omega_6\}\right) = 0.05,$$

נגדיר את המאורעות הבאים: $E = \{\omega_1, \omega_2\}$, $F = \{\omega_2, \omega_3\omega_4\}$ ו־ $G = \{\omega_2, \omega_3, \omega_5\}$.

(א) חשבו את ההסתברויות של המאורעות הבאים: $\overline{E}$, $\overline{F}$, $E \cup \overline{F}$, $E \cup G$ ו־ $E \cup F \cup G$.

(ב) בדקו האם מתקיים $\mathrm{P}(E \cup F) = \mathrm{P}(E) + \mathrm{P}(F)$, הסבירו.

(ג) הוכיחו כי P כפי שהוגדרה היא פונקצית הסתברות.

6. היזכרו בתכונות של פונקצית ההסתברות וענו על השאלות הבאות:

(א) נתונים זוג מאורעות **זרים** $\Omega \subseteq C$ ו־$\Omega \subseteq D$ המקיימים: $\mathrm{P}\left(\overline{C} \cap D\right) = \mathrm{P}\left(C \cap \overline{D}\right)$ ו־ $\mathrm{P}(C) = 0.2$. חשבו את $\mathrm{P}(C \cup D)$.

(ב) יהיו A ו־B שני מאורעות במרחב המדגם. ידוע ש־ $\mathrm{P}\left(B\right) = 0.7$ וכי $\mathrm{P}\left(A\right) = 0.4$. הוכיחו כי מתקיים: $0.1 \leq \mathrm{P}(A \cap B) \leq 0.4$.

(ג) נתונים זוג מאורעות $\Omega \subseteq A$ ו־$\Omega \subseteq B$ המקיימים: $\mathrm{P}\left(\overline{A} \cap \overline{B}\right) = 0.1$, $\mathrm{P}\left(\overline{B}\right) = 0.2$ ו־ $\mathrm{P}\left(A \cap B\right) = 0.6$. חשבו את $\mathrm{P}(A)$.

(ד) ראיתם שלכל זוג מאורעות $A_1, A_2 \subseteq \Omega$ מתקיים:

$$\mathrm{P}\left(A_1 \cup A_2\right) = \mathrm{P}\left(A_1\right) + \mathrm{P}\left(A_2\right) - \mathrm{P}\left(A_1 \cap A_2\right)$$

השתמשו בכך בכדי להראות שלכל שלשה של מאורעות $A_1, A_2, A_3 \subseteq \Omega$ מתקיים:

$$\mathrm{P}\left(A_1 \cup A_2 \cup A_3\right) =$$

$$\mathrm{P}\left(A_1\right) + \mathrm{P}\left(A_2\right) + \mathrm{P}\left(A_3\right) - \mathrm{P}\left(A_1 \cap A_2\right) - \mathrm{P}\left(A_1 \cap A_3\right) - \mathrm{P}\left(A_2 \cap A_3\right) + \mathrm{P}\left(A_1 \cap A_2 \cap A_3\right)$$

7. בהכנת תמיסה רפואית, עלולות להתרחש תקלות משני סוגים: תקלת `ריכוז` (ריכוז לא נכון) שהסתברותה 0.1, ותקלת `זיהום` (ע`י גורמים שלא אמורים להיות בתמיסה) שהסתברותה 0.04. ההסתברות שלפחות תקלה אחת תתרחש היא 0.13

(א) הגדירו את מרחב המדגם של הניסוי

(ב) מהי ההסתברות ששתי תקלות יתרחשו בו זמנית?

(ג) מהי ההסתברות שלכל היותר תקלה אחת תתרחש?

(ד) מהי ההסתברות שאף תקלה לא תתרחש?

חלק III
קומבינטוריקה

11 מדגמים

לניסוי פשוט יש n תוצאות אפשריות.]למשל בהטלת קוביה יש 6 תוצאות אפשריות[. נניח שחוזרים על הניסוי הפשוט r פעמים.]למשל מטילים קוביה פעמיים[. השאלה המרכזית שנרצה לברר במסגרת זו היא כמה תוצאות קיימות לניסוי המורכב? במילים אחרות, מהי עוצמתו של מרחב המדגם?

ראשית, בכל פעם שמבצעים ניסוי חשוב להבחין בשני מאפיינים:

- האם יש או אין חשיבות ל**סדר**?
- האם הניסוי מתבצע עם או בלי **החזרה**?

למשל בהטלת קוביה פעמיים ניתן לקבל תוצאה של $(1,2)$ או תוצאה של $(2,1)$. עלינו להחליט לפי המקרה האם מדובר בשתי תוצאות שונות או בתוצאה אחת. כלומר, האם יש חשיבות לסדר או לא. למשל בשליפת פתק מתוך כובע פעמיים יש חשיבות לשאלה האם לפני השליפה השנייה אנו מחזירים את הפתק שיצא בשליפה הראשונה או לא. כלומר, האם מדובר במדגם עם החזרה או בלי החזרה.

להלן נדון בכל האפשרויות: מדגם סדור/לא סדור עם החזרה/בלי החזרה.

11.1 מדגם סדור עם החזרה

במדגם סדור עם החזרה מספר האפשרויות הוא $|\Omega| = n^r$.

למשל בהטלת קוביה פעמיים, מספר האפשרויות הוא $6^2 = 36$:

$$\Omega = \begin{bmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{bmatrix} \Rightarrow |\Omega| = 6^2 = 36$$

נשים לב שמכיוון שיש חשיבות לסדר, מנינו גם את האפשרות $(1,3)$ ו–$(3,1)$ כשתי אפשרויות שונות.

11.2 מדגם סדור ללא החזרה

הגדרה: `עצרת` של מספר טבעי k מוגדרת ומסומנת כך:

$$k! = 1 \cdot 2 \cdot 3 \cdot ... \cdot k$$

במדגם סדור ללא החזרה מספר האפשרויות הוא $|\Omega| = \frac{n!}{(n-r)!}$, $0 \le r \le n$

בדוגמה של הטלת הקוביה, העובדה שלא מאפשרים חזרה מסירה את כל האפשרויות מהאלכסון בו מוצגות התוצאות בהן שתי התוצאות זהות. לכן נישאר עם האפשרויות הבאות:

$$\Omega = \begin{bmatrix} & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & \end{bmatrix} \Rightarrow |\Omega| = \frac{6!}{(6-2)!} = 6 \cdot 5 = 30$$

נסביר כיצד הגענו לנוסחה: בניסוי הפשוט הראשון קיימות n תוצאות אפשריות. בעקבות הניסוי הראשון ואין זה משנה מה היתה תוצאתו. בניסוי הפשוט השני ירדה אפשרות אחת (כי אין החזרה) ולכן נשארנו עם $n-1$ תוצאות אפשריות, וכן הלאה. בניסוי הפשוט ה־r נישאר עם $n-r+1$ תוצאות אפשריות.
מכאן שסך האפשרויות הוא:

$$n \cdot (n-1) \cdot \ldots \cdot (n-r+1) =$$

$$\frac{n \cdot (n-1) \cdot \ldots \cdot (n-r+1) \cdot (n-r) \cdot \ldots \cdot 2 \cdot 1}{(n-r) \cdot \ldots \cdot 2 \cdot 1} = \frac{n!}{(n-r)!}$$

לצורך שלמות ההגדרה, נאמר כי $0! = 1$. נראה בהמשך שהגדרה זו שימושית במקרי־קיצון. כך למשל לפי הגדרה זו מספר התוצאות האפשריות של סידור ללא החזרה של n איברים הוא $\frac{n!}{(n-n)!} = n!$.

11.3 מדגם לא סדור ללא החזרה

בהשוואה למדגם סדור עם החזרה, מספר האפשרויות מצטמצם, כי מאורעות בעלי אותם איברים בסדר שונה מתלכדים למאורע אחד.

במדגם לא סדור ללא החזרה מספר האפשרויות הוא $|\Omega| = \binom{n}{r} = \frac{n!}{r!(n-r)!}$,
$0 \le r \le n$

$\binom{n}{r}$הוא סימון לביטוי האלגברי $\frac{n!}{r!(n-r)!}$וקוראים אותו `n על r"

מספר זה נותן למעשה את מספר הצירופים האפשריים של r איברים מתוך n איברים.
נסביר כיצד הגענו לנוסחה: ראשית נתונות לנו $\frac{n!}{(n-r)!}$ אפשרויות בהנחה שהסדר משנה. כעת נרצה להסיר האפשרויות שמופיעות יותר מפעם אחת ולמנות אותן רק פעם אחת. נשים לב כי ראינו שעבור כל r איברים נתונים קיימים $r!$ סידורים שונים אפשריים, כי במקום הראשון יש r אפשרויות, במקום השני $r-1$ אפשרויות וכן הלאה. לכן נחלק ב־$r!$ ונקבל

את הנוסחה שקבענו.

נשים לב שבמקרים $r = 0$, $r = n$ נקבל: $\binom{n}{0} = \binom{n}{n} = 1$ (נזכור שהגדרנו $0! = 1$). ההיגיון בתוצאה זו הוא שמתוך n איברים יש רק דרך אחת לבחור 0 איברים או n איברים ללא חשיבות לסדר.

טענה:

$$\binom{n}{r} = \binom{n}{n-r}$$

הוכחה:

$$\frac{n!}{(n-r)!\,(n-(n-r))!} = \frac{n!}{(n-r)!r!} = \frac{n!}{r!\,(n-r)!}$$

■

ההסבר לשוויון זה הוא שמדובר באירועים משלימים אחד לשני. למשל קל לראות שבחירת 3 תלמידים מכיתה של 10 לחברות בוועד, זו פעולה שקולה לבחירת 7 תלמידים מכיתה של 10 שלא יהיו חברים בוועד.

טענה:

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$$

הוכחה:

$$\binom{n-1}{r-1} + \binom{n-1}{r} = \frac{(n-1)!}{(r-1)!\,((n-1)-(r-1))!} + \frac{(n-1)!}{r!\,(n-r-1)!}$$

$$= \frac{(n-1)!}{(r-1)!\,(n-r)!} + \frac{(n-1)!}{r!\,(n-r-1)!} = \frac{r\,(n-1)!}{r!\,(n-r)!} + \frac{(n-r)\,(n-1)!}{r!\,(n-r)!}$$

$$= \frac{r\,(n-1)! + (n-r)\,(n-1)!}{r!\,(n-r)!} = \frac{(r+n-r)\,(n-1)!}{r!\,(n-r)!}$$

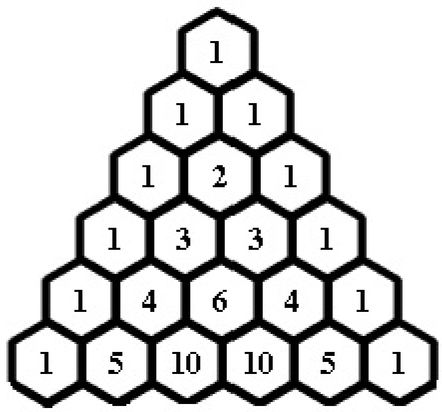
$$= \frac{n\,(n-1)!}{r!\,(n-r)!} = \frac{n!}{r!\,(n-r)!} = \binom{n}{r}$$

■

נסביר את השוויון שקיבלנו. נניח שבוחרים r איברים מתוך n, ונניח ש־x הוא איבר כלשהו מתוך ה־n. ברור שיש שתי אפשרויות זרות: או ש־x כלול ב־r האיברים הנבחרים או שלא. אין אפשרות נוספת. מכיוון שהאפשרויות הללו זרות, אם נחשב את מספר התוצאות האפשריות בכל אחת מהאפשרויות ונסכום, נקבל את כל התוצאות האפשריות. במקרה ש־x כלול ב־r שבחרנו, נשאר לבחור $r-1$ עצמים מתוך $n-1$

העצמים הנותרים, ולכן מספר התוצאות האפשריות הוא $\begin{pmatrix} n-1 \\ r-1 \end{pmatrix}$. במקרה ש-$x$ לא כלול ב-r שבחרנו, עלינו לבחור r איברים מתוך $n-1$ האיברים הנותרים (כי את x כבר אי־אפשר לבחור), ולכן מספר התוצאות האפשריות הוא $\begin{pmatrix} n-1 \\ r \end{pmatrix}$.

באמצעות הטענה האחרונה ניתן להציג את ערכו של $\begin{pmatrix} n \\ r \end{pmatrix}$ באמצעות מה שמכונה `משולש פסקל`:



המשולש נבנה כך שכל ערך מתקבל כסכום שני הערכים שמעליו. זו בדיוק הטענה שהוכחנו כעת ולכן הערך ה-r בשורה ב-n הוא $\begin{pmatrix} n \\ r \end{pmatrix}$.

11.3.1 הבינום של ניוטון

$$(a+b)^n = \sum_{r=0}^{n} \begin{pmatrix} n \\ r \end{pmatrix} a^r b^{n-r}$$

ראשית ברור שכאשר a נכפל r פעמים, נשאר ל-b להיות נכפל $n-r$ פעמים. כדי לבדוק את כל הקומבינציות האפשריות שמתקבלות, משתמשים במה שהוכחנו לעיל אודות בחירת r איברים מתוך n, כאשר הסדר לא משנה וללא החזרה.

כך זה נראה במקרה של $n=4$:

$$(a+b)^4=(a+b)(a+b)(a+b)(a+b)$$

$$=\binom{4}{4}b^4+\binom{4}{3}ab^3+\binom{4}{2}a^2b^2+\binom{4}{1}a^3b+\binom{4}{0}a^4b^4$$

$$=\sum_{r=0}^{4}\binom{4}{r}a^rb^{4-r}$$

כאשר השיוויון השני מקורו במספר האפשרויות שלנו לבחור את הסוגריים מהן יבוא הגורם a (ומכל השאר הגורם b).

הרחבה: אם $a=b=1$ מתקיים:

$$2^n=(1+1)^n=\sum_{r=0}^{n}\binom{n}{r}1^r1^{n-r}=\sum_{r=0}^{n}\binom{n}{r}=\binom{n}{0}+\binom{n}{1}+...+\binom{n}{n}$$

כלומר 2^n זה מספר סך המאורעות שניתן ליצור במרחב מדגם המכיל n איברים.
הסיבה לכך היא שניתן ליצור $\binom{n}{0}$ מאורעות בעלי 0 מאורעות פשוטים, $\binom{n}{1}$ מאורעות בעלי 1 מאורעות פשוטים, וכן הלאה.

11.4 מדגם לא סדור עם החזרה

במדגם לא סדור עם החזרה מספר האפשרויות הוא
$|\Omega|=\binom{n+r-1}{r}=\frac{(n+r-1)!}{r!\,(n-1)!}$.

לא נוכיח תוצאה זו אולם נדגים אותה.
במשחק שש־בש זורקים שתי קוביות בבת־אחת. לצורך מהלך המשחק אין הבדל למשל בין $(2,3)$ לבין $(3,2)$ ולכן הסדר לא משנה. התוצאות האפשריות הן:

$$\Omega=\begin{bmatrix}(1,1)&&&&&\\(2,1)&(2,2)&&&&\\(3,1)&(3,2)&(3,3)&&&\\(4,1)&(4,2)&(4,3)&(4,4)&&\\(5,1)&(5,2)&(5,3)&(5,4)&(5,5)&\\(6,1)&(6,2)&(6,3)&(6,4)&(6,5)&(6,6)\end{bmatrix}$$

כלומר $n=6$, $r=2$, ולכן $|\Omega|=\binom{7}{2}=21$.

הערה: במקרה זה אין מקום להניח הסתברויות אחידות מעל 21 האפשרויות הנ"ל, כי מכיוון שהסדר לא משנה, ספרנו למשל את $(1,2)$ פעם אחת, בעוד התוצאה הזו מתקבלת גם על־ידי $(2,1)$ ולכן סביר לתת לה הסתברות כפולה.

11.5 דוגמאות למרחב הסתברות אחיד

11.5.1 זריקת קוביות

$$\Omega = \begin{bmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{bmatrix}$$

$|\Omega| = n^r = 6^2 = 36$. נניח פונקצית הסתברות אחידה, ולכן לכל תוצאה הסתברות $\frac{1}{36}$.

- ההסתברות ל`3 בקוביה הראשונה`$=A$:

$$\mathrm{P}\left(A\right) = \frac{6}{36} = \frac{1}{6}$$

- ההסתברות ל 3 בזריקה השנייה`$=B$:

$$\mathrm{P}\left(A\right) = \frac{6}{36} = \frac{1}{6}$$

- ההסתברות ל`לפחות פעם אחת 3 בשתי הזריקות`$=A \cup B$:

$$\mathrm{P}\left(A \cup B\right) = \frac{6+5}{36} = \frac{11}{36}$$

נשים לב שהחיתוך $A \cap B$ מכיל את $(3,3)$ ולכן $P\left(A \cap B\right) = \frac{1}{36}$, ומכאן נקבל את אותה התוצאה באופן הבא:

$$\mathrm{P}\left(A \cup B\right) = \mathrm{P}\left(A\right) + \mathrm{P}\left(B\right) - \mathrm{P}\left(A \cap B\right) = \frac{6}{36} + \frac{6}{36} - \frac{1}{36} = \frac{11}{36}$$

- ההסתברות ל`לא לקבל 3 `$=\overline{A \cup B}$:

$$\mathrm{P}\left(\overline{A \cup B}\right) = 1 - \mathrm{P}\left(A \cup B\right) = 1 - \frac{11}{36} = \frac{25}{36}$$

11.5.2 ימי־הולדת

נבדוק את ההסתברות שבכיתה בת 35 סטודנטים, כולם חוגגים יום־הולדת בימים שונים במהלך השנה, תחת ההנחה שכל הימים שווי־הסתברות.
$n = 365$, $r = 35$ ולכן $|\Omega| = 365^{35}$. לצורך עמידה בתנאים, לסטודנט הראשון 365 אפשרויות, לשני 364, לשלישי 363, וכן הלאה, עד שלסטודנט ה־35 נותרו $365 - 35 + 1 =$

331$.$ נשים לב שזהו מדגם סדור וללא החזרה, ולכן מספר האפשרויות לקומבינציה של 35 ימי הולדת בימים שונים, הוא $\frac{n!}{(n-r)!} = \frac{365!}{330!}$. סך כל האפשרויות לימי־הולדת הוא 365^{35}, ולכן ההסתברות לאירוע המבוקש היא:

$$\frac{\frac{365!}{330!}}{365^{35}} = 0.17$$

האם ציפית להסתברות גבוהה יותר? אגב, בכיתה בת 23 תלמידים ההסתברות היא חצי לערך.

11.5.3 זריקת כדורים לתאים

נניח שזורקים שלושה כדורים לשלושה תאים, כאשר כל תא יכול להכיל את כל שלושת הכדורים. מספר האפשרויות לסידור הכדורים בתאים הוא $|\Omega| = 3^3 = 27$.

- ההסתברות שכל התאים יהיו מלאים היא:

$$\frac{3!}{3^3} = \frac{6}{27} = \frac{2}{9}$$

כי מספר הדרכים לסדר שלושה כדורים בשלושה תאים שונים הוא $3!$.

- ההסתברות שכל הכדורים יהיו באותו התא היא:

$$\frac{3}{3^3} = \frac{1}{9}$$

כי מספר הדרכים לסדר את כל שלושת הכדורים באותו כד, הוא 3.

- ההסתברות שתא אחד בדיוק יישאר ריק היא $\frac{2}{3}$. הסבר: ההסתברות שתא אחד בדיוק יישאר ריק היא ההסתברות המשלימה של המקרה בו יש שני תאים ריקים ושל המקרה בו כל התאים מלאים. לעיל מצאנו שההסתברויות הללו הן $\frac{1}{9}$ ו־$\frac{2}{9}$, בהתאמה, ולכן ההסתברות המבוקשת היא:

$$1 - \frac{1}{9} - \frac{2}{9} = \frac{6}{9} = \frac{2}{3}$$

דרך נוספת היא לחשב זאת באופן ישיר. ראשית נקבע תא ריק (יש 3 אפשרויות לכך). את שלושת הכדורים ניתן לפזר ב־$8 = 2^3$ דרכים שונות בין שני התאים האחרים. נסיר שתי אפשרויות שבהן יש תא ריק נוסף (אם כל הכדורים בתא אחד), ונישאר עם 6 אפשרויות. לכן יש לנו $3 \cdot 6 = 18$ אפשרויות, וההסתברות המבוקשת היא:

$$\frac{18}{27} = \frac{2}{3}$$

11.5.4 קלפי ברידג'

שחקן ברידג, מקבל 13 קלפים מתוך חפיסה של 52.

- השחקן מעוניין לדעת מהי ההסתברות שמתוך 13 הקלפים יהיו לו 5 הקלפים הבאים: אס, מלך, מלכה, נסיך ו–10־עלה. ראשית מספר כל האפשרויות לקבל 13 מתוך 52 הוא $\binom{52}{13}$. השחקן מעוניין ב–5 קלפים מסוימים, ולא משנה לו מה יהיו שאר 8 הקלפים. אירוע המטרה שלנו הוא לבחור את 5 הקלפים המבוקשים, ובהינתן שבחרנו אותם נותרו $\binom{47}{8}$ אפשרויות לבחור את שאר הקלפים. מכאן שההסתברות למאורע המבוקש היא:

$$\frac{\binom{47}{8}}{\binom{52}{13}}$$

נשים לב שגם לו היינו בוחרים 5 קלפים מסוימים אחרים ההסתברות לא הייתה שונה.

- כעת נניח שהשחקן מעוניין לקבל 5 קלפים מסוימים, כולם בצורת לב או 5 קלפים מסוימים כולם בצורת עלה. נזכור את הנוסחה: $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$. מכאן שההסתברות למאורע המבוקש, שהוא איחוד של מאורעות, היא:

$$\frac{2 \cdot \binom{47}{8} - \binom{42}{3}}{\binom{52}{13}}$$

כי $\binom{42}{3}$ זה מספר האפשרויות לבחור את 3 הקלפים האחרים מתוך 42 הקלפים הנותרים.

11.5.5 חברי־כנסת

נניח שבכנסת 20 חברים ממפלגה א' ו–30 חברים ממפלגה ב'. בוחרים מתוכם באקראי 2 חברי־כנסת.

- ההסתברות ששניהם ממפלגה א' היא:

$$\frac{\binom{20}{2}}{\binom{50}{2}}$$

- ההסתברות ששניהם ממפלגה ב' היא:

$$\frac{\binom{30}{2}}{\binom{50}{2}}$$

- ההסתברות שאחד ממפלגה א' והאחר ממפלגה ב' היא:

$$\frac{\binom{20}{1}\cdot\binom{30}{1}}{\binom{50}{2}}$$

11.6 הסתברויות היפר־גאומטריות

נניח שנתונה אוכלוסייה בגודל N, כאשר k מתוכם שייכים לסוג א' והשאר $(N-k)$ שייכים לסוג ב'. בוחרים מתוך האוכלוסייה מדגם בגודל r, ההסתברות ש־x מתוך המדגם הם מסוג א' היא:

$$\frac{\binom{k}{x}\cdot\binom{N-k}{r-x}}{\binom{N}{r}}$$

כאשר $\max\{0,r+k-N\}\le x\le\min\{k,r\}$.

עבור כל ערכי x שאינם בתחום זה, ההסתברות היא אפס. הדרישה $x\le\min\{k,r\}$ ברורה. מקור הדרישה $\max\{0,r+k-N\}\le x$ היא בכך שמספר הנבחרים מסוג ב', $r-x$, אינו יכול לעלות על מספרם באוכלוסיה $N-k$.

12 שאלות חזרה

1. כמה סיסמאות ניתן ליצור מהאותיות `סמבטיון` כאשר נדרוש סיסמאות עד אורך של שבע אותיות וניתן להשתמש בכל אות יותר מפעם אחת? (לדוגמא הסיסמאות: ס, סס, ממססב, הן חוקיות)

2. כמה סיסמאות ניתן ליצור מהאותיות `אנדרלמוסיה` כאשר נדרוש סיסמאות עד אורך של שבע אותיות וניתן להשתמש בכל אות פעם אחת בדיוק?

3. על המדף מונחים 22 תבלינים שונים. בכמה דרכים ניתן לתבל אורז לבן ב־3 תבלינים שונים?

4. מתוך כיתה של 10 תלמידים ו־7 תלמידות, בוחרים ועד המורכב מ־ 5 תלמידים ו־3 תלמידות. כמה אפשרויות קיימות להרכב הועד הנבחר? (אין חשיבות לסדר בתוך הועד) מה ההסתברות שמיכל ונדב (מועמדים לועד) יבחרו?

5. בכמה אופנים ניתן לסדר קבוצה של 7 ספרי מתמטיקה שונים, ספרי פיזיקה שונים ו־ 3 ספרי כימיה שונים על מדף? מה תהיה התשובה אם הספרים בכל נושא חייבים להיות סמוכים זה לזה?

6. בכמה אופנים שונים אפשר לסדר ארבעה ספרים שונים בשבעה תאים אם בכל תא יש מקום לספר אחד? ומה אם אין הגבלה כזו?

7. במרוץ מתחרות שמונה אצניות. מישראל, ירדן, מצרים, עירק, איראן, סוריה, לבנון וטורקיה. האצנית המצרית לא מוכנה לרוץ ליד האצנית הלבנונית. כמה אפשרויות יש לסדר את האצניות באופן כזה? מה ההסתברות כי הטורקיה והלבונונית יוגרלו להיות זו לצד זו?

8. בבניין משותף יש ארבע כניסות, ובכל כניסה עשר דירות. בועד של שישה אנשים חשוב שיהיה יצוג לכל כניסה, ומכל דירה יכול להשתתף בועד אדם אחד. לכל אדם בועד תפקיד שונה, בכמה צורות ניתן להרכיב את הועד?

9. עשרה תלמידים עולים על הסעה מתחנה קבועה. במסלול חמישה בתי ספר וכל תלמיד יורד מההסעה בבית הספר שלו. הנהג רושם כמה נוסעים יורדים בכל תחנה, כמה אפשרויות שונות יש?

10. ליגה היא אוסף של קבוצות. בשלב הראשון של התחרות מוגרלות זוגות של קבוצות שמתחרות בינהן. אם בליגה יש עשר קבוצות, כמה אפשרויות שונות יש לסידור קבוצות של השלב הראשון?

חלק IV

הסתברות מותנה (Conditional probability) ואי־תלות (Independence)

13 הסתברות מותנה ובאיסניות

הסתברות מותנה היא מידת הוודאות למאורע, בהינתן שקרתה עובדה כלשהי שרלוונטית למאורע. למשל, ההסתברות שברק אובמה נבחר לנשיאות ארצות הברית תלויה בשאלה מי זכה בבחירות המקדימות במפלגה הדמוקרטית. אם היינו יודעים שהילרי קלינטון זכתה ההסתברות היא 0, ואם היינו יודעים שאובמה זכה קיימת הסתברות חיובית.

הגדרה: נניח כי A,B מאורעות במרחב מדגם $|\Omega|$, ונניח $\mathrm{P}\left(B\right)>0$.
ההסתברות של A בהינתן B שנסמן $\mathrm{P}\left(A/B\right)$ מוגדרת:

$$\mathrm{P}\left(A|B\right)=\frac{\mathrm{P}\left(A\cap B\right)}{\mathrm{P}\left(B\right)}$$

נשים לב שבהעברת אגפים פשוטה מקבלים את השוויון:

$$\mathrm{P}\left(A\cap B\right)=\mathrm{P}\left(A|B\right)P\left(B\right)$$

במקרה של מרחב הסתברות אחיד נקבל כי $\mathrm{P}(A|B)=\frac{|A\cap B|}{|B|}$

דוגמה 1: מטילים זוג קוביות הוגנות. נגדיר את המאורעות הבאים: `הסכום על פני הקוביות הוא 9` =A. מתקיים כי $\mathrm{P}(A)=\frac{4}{36}=\frac{1}{9}$. `בקוביה השנייה מתקבל 5` = B. מתקיים כי $\mathrm{P}(B)=\frac{1}{6}$. כמו־כן, $\mathrm{P}(A\cap B)=\frac{1}{36}$.

כעת נניח שקיבלנו מידע שבקוביה השנייה התקבל 5. מהי ההסתברות כעת שהסכום על פני שתי הקוביות הוא 9? נשתמש בהגדרה של הסתברות מותנה:

$$\mathrm{P}(A|B)=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)}=\frac{\frac{1}{36}}{\frac{1}{6}}=\frac{6}{36}=\frac{1}{6}$$

מכאן שתוספת המידע אודות התוצאה של הקוביה השנייה שינתה (במקרה זה הגדילה מ־$\frac{1}{9}$ל־$\frac{1}{6}$) את ההסתברות למאורע שהגדרנו.
ניתן גם לבדוק את ההפך: מהי ההסתברות שבקוביה השנייה קיבלנו 5, אם נתון שסכום הקוביות הוא 9:

$$\mathrm{P}(B|A)=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(A)}=\frac{\frac{1}{36}}{\frac{1}{9}}=\frac{9}{36}=\frac{1}{4}$$

וגם כאן הסתברותו של המאורע עלתה בעקבות המידע החדש (מ־$\frac{1}{6}$ל־$\frac{1}{4}$).

תכונות ההסתברות המותנה: כל התכונות של הסתברות נשמרות גם להסתברות מותנה. כך למשל:

$$\mathrm{P}(\phi|B)=0$$

$$\mathrm{P}(\Omega|B)=1$$

$$0\leq \mathrm{P}(A|B)\leq 1$$

$$\mathrm{P}(A\cup C|B)=\mathrm{P}(A|B)+\mathrm{P}(C|B)-\mathrm{P}(A\cap C|B)$$

כמו כן נניח שנתון $\Omega=\{\omega_1,\omega_2,...,\omega_n\}$ ומוגדר המאורע $A=\{\omega_1,\omega_2\}$. יחס ההסתברויות בין המאורעות הפשוטים ω_1,ω_2 נשמר גם בהינתן ש־A התרחש:

$$\frac{\mathrm{P}(\omega_1|A)}{\mathrm{P}(\omega_2|A)}=\frac{\frac{\mathrm{P}(\omega_1\cap A)}{\mathrm{P}(A)}}{\frac{\mathrm{P}(\omega_2\cap A)}{\mathrm{P}(A)}}=\frac{\mathrm{P}(\omega_1)}{\mathrm{P}(\omega_2)}$$

כלומר, פונקציית ההסתברות המותנה במאורע משמרת את היחסים הפנימיים בין הסתברויות המאורעות הפשוטים המרכיבים אותן.

הרחבה: נזכור את הנוסחה $\mathrm{P}(A)=\mathrm{P}(A\cap B)+\mathrm{P}(A\cap \overline{B})$.
ומהגדרת הסתברות מותנה נובע כי:

$$\boxed{\mathrm{P}(A)=\mathrm{P}(B)\cdot\mathrm{P}(A|B)+\mathrm{P}\left(\overline{B}\right)\cdot\mathrm{P}\left(A|\overline{B}\right)}$$

המשמעות של שוויון זה היא שההסתברות $\mathrm{P}(A)$ היא שקלול בין $\mathrm{P}(A|B)$ לבין $\mathrm{P}\left(A|\overline{B}\right)$, כאשר המשקל של כל אחד מהם הוא $P(B)$ ו-$\mathrm{P}\left(\overline{B}\right)$ בהתאמה. לכן אם $\mathrm{P}(A|B) > \mathrm{P}(A)$ (כלומר התרחשות המאורע B הגדילה את ההסתברות ל-A) אז בהכרח $\mathrm{P}\left(A|\overline{B}\right) < \mathrm{P}(A)$. ז`א התרחשות משלים של B הורידה את ההסתברות של A.

13.1 נוסחת ההסתברות השלמה

נניח ש-$\{B_i\}_{i=1}^{n}$ היא חלוקה של מרחב המדגם. כלומר לכל $i \neq j$ מתקיים $B_i \cap B_j = \phi$, וכן $\bigcup_{i=1}^{n} B_i = \Omega$. אז מתקיים:

$$\boxed{\mathrm{P}(A) = \sum_{i=1}^{n} \mathrm{P}(A \cap B_i) = \sum_{i=1}^{n} \mathrm{P}(B_i)\,\mathrm{P}(A|B_i)}$$

נוסחה זו שימושית כאשר נתונות הסתברויות של מאורע המותנית במאורעות אחרים, כאשר המאורעות האחרים יוצרים חלוקה.

דוגמה: נתונים 3 כדים המכילים כדורים.

- בכד a כדור לבן וכדור אדום
- בכד b כדור לבן ושני כדורים אדומים
- בכד c כדור לבן ושלושה כדורים אדומים

נבחר כד באופן מקרי (כך שההסתברות לכל כד היא $\frac{1}{3}$) ונבחר ממנו כדור באופן מקרי.

- מה ההסתברות לכדור לבן?

$$\mathrm{P}(\text{white}|a) = \tfrac{1}{2}$$

$$\mathrm{P}(\text{white}|b) = \tfrac{1}{3}$$

$$\mathrm{P}(\text{white}|c) = \tfrac{1}{4}$$

נסכום את ההסתברויות באופן משוקלל:

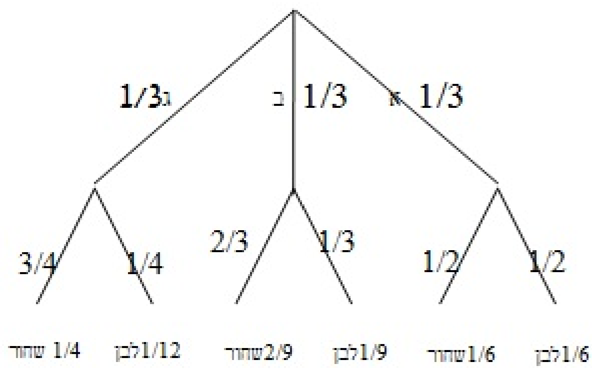
$$\mathrm{P}(\text{white}) = \tfrac{1}{3} \cdot \mathrm{P}(\text{white}|a) + \tfrac{1}{3} \cdot \mathrm{P}(\text{white}|b) + \tfrac{1}{3} \cdot \mathrm{P}(\text{white}|c) =$$

$$= \tfrac{1}{3} \cdot \tfrac{1}{2} + \tfrac{1}{3} \cdot \tfrac{1}{3} + \tfrac{1}{3} \cdot \tfrac{1}{4} = \tfrac{13}{36}$$

ההסתברות לכדור שחור היא כמובן ההסתברות המשלימה:

$$\mathrm{P}(\text{black}) = 1 - \frac{13}{36} = \frac{23}{36}$$

שאלות מסוג זה ניתנות להצגה באמצעות `**עץ**`.
כך למשל הדוגמה האחרונה מוצגת באמצעות העץ הבא:



ההסתברות שמופיעה לכל אירוע בכל קצה של העץ, היא מכפלת ההסתברויות לאורך הענפים המובילים אליו.
דרך נוספת להציג את הדוגמה שהזכרנו:

צבע/	שחור	לבן	סכום
כד a	$^1/_6$	$^1/_6$	$^1/_3$
כד b	$^2/_9$	$^1/_9$	$^1/_3$
כד c	$^1/_4$	$^1/_{12}$	$/_3$
סכום	$^{23}/_{36}$	$^{13}/_{36}$	1

13.2 נוסחת ביאס (Bayes' theorem)

בהינתן חלוקה $\{B_i\}_{i=1}^n$ של מרחב מדגם Ω ומאורע כלשהו A, לכל $1 \leq i \leq n$ מתקיים:

$$\mathrm{P}\left(B_i|A\right) = \frac{\mathrm{P}\left(A \cap B_i\right)}{\mathrm{P}\left(A\right)} = \frac{\mathrm{P}\left(B_i\right)\mathrm{P}\left(A|B_i\right)}{\sum_{j=1}^n \mathrm{P}\left(A \cap B_j\right)} = \frac{\mathrm{P}\left(B_i\right)\mathrm{P}\left(A|B_i\right)}{\sum_{j=1}^n \mathrm{P}\left(B_j\right)\mathrm{P}\left(A|B_j\right)}$$

נוסחה זו נובעת מהגדרת הסתברות מותנית ומנוסחת ההסתברות השלמה. נשים לב שביצענו היפוך של התנאי. במקום לדון בהסתברות של מאורע המטרה A בהינתן B_i, אנו מחשבים את הסתברות B_i בהינתן A. כמובן ניתן לחשב הסתברות זו רק בהינתן ההסתברויות $\mathrm{P}\left(B_j\right)$, $1 \leq j \leq n$.

המשך הדיון בדוגמה

נרצה להפוך את הדיון. כלומר נניח שידוע שיצא כדור לבן, ונשאל מה ההסתברות שהכדור הגיע מכל אחד מהכדים?

נחשב:

$$\mathrm{P}\left(a|\mathrm{white}\right)=\frac{\mathrm{P}\left(a\cap\mathrm{white}\right)}{\mathrm{P}\left(\mathrm{white}\right)}=\frac{\mathrm{P}\left(a\right)P\left(\mathrm{white}|a\right)}{\mathrm{P}\left(\mathrm{white}\right)}=\frac{\frac{1}{6}}{\frac{13}{36}}=\frac{6}{13}>\frac{1}{3}=P\left(a\right)$$

$$\mathrm{P}\left(b|white\right)=\frac{\mathrm{P}\left(b\cap\mathrm{white}\right)}{\mathrm{P}\left(\mathrm{white}\right)}=\frac{\mathrm{P}\left(b\right)\mathrm{P}\left(\mathrm{white}|b\right)}{\mathrm{P}\left(\mathrm{white}\right)}=\frac{\frac{1}{9}}{\frac{13}{36}}=\frac{4}{13}<\frac{1}{3}=P\left(b\right)$$

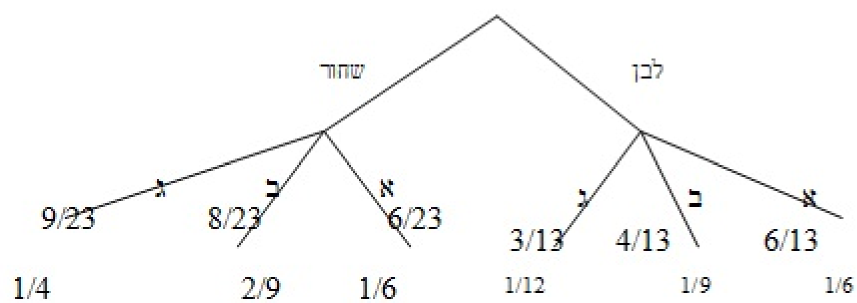
$$\mathrm{P}\left(c|white\right)=\frac{\mathrm{P}\left(c\cap\mathrm{white}\right)}{\mathrm{P}\left(\mathrm{white}\right)}=\frac{\mathrm{P}\left(c\right)\mathrm{P}\left(\mathrm{white}|c\right)}{\mathrm{P}\left(\mathrm{white}\right)}=\frac{\frac{1}{12}}{\frac{13}{36}}=\frac{3}{13}<\frac{1}{3}=P\left(c\right)$$

דרך נוספת לחשב את $P\left(c|\mathrm{white}\right)$ לאחר שידועות שתי ההסתברויות האחרות, היא:

$$\mathrm{P}\left(c|white\right)=1-\mathrm{P}\left(a|\mathrm{white}\right)-\mathrm{P}\left(b|\mathrm{white}\right)=1-\frac{6}{13}-\frac{4}{13}=\frac{3}{13}$$

באופן אינטואיטיבי, קל היה לנחש ש־$\mathrm{P}\left(a/\mathrm{white}\right)>\frac{1}{3}$ וכן $\mathrm{P}\left(c/\mathrm{white}\right)<\frac{1}{3}$, שכן כל אחד מהם הוא מקרה קצה של מספר לבנים, אבל המקרה $\mathrm{P}\left(b/white\right)$ קשה יותר לניחוש.

ניתן לבנות עץ אפשרויות גם להיפוך, הסדר:



במבט ראשון עץ זה נראה מלאכותי בהתחשב בדינמיקה שהגדירה את הניסוי. אבל מצד שני מוביל אותנו לכך שמה שחשוב היא האינפורמציה החלקית על הניסוי שנמסרה ולא קשר כלשהו שלה לזמן.

דוגמה: נניח שבכד 220 כדורים, מתוכם 20 אדומים ו־200 לבנים. מוציאים כדור באופן אקראי. אם הכדור שהוצא אדום, הוא מוחזר לכד ומוסיפים לכד 3 כדורים אדומים. אם הכדור שהוצא לבן, הוא מוחזר לכד ומוסיפים לכד 2 כדורים לבנים. לבסוף, מוציאים כדור נוסף אקראי מהכד. נסמן ב־A_i את המאורע של הוצאת כדור אדום בפעם ה־i, כאשר $1\leq i\leq 2$. נסו להעריך מראש מי גדול ממי – $\mathrm{P}\left(A_1\right)$ או $\mathrm{P}\left(A_2\right)$.

נחשב באופן מדויק:

$$\mathrm{P}\left(A_1\right)=\tfrac{20}{220}=\tfrac{1}{11}$$

$$\begin{aligned}\mathrm{P}\left(A_2\right)&=\mathrm{P}\left(A_2\cap A_1\right)+\mathrm{P}\left(A_2\cap\overline{A_1}\right)\\&=\mathrm{P}\left(A_1\right)\mathrm{P}\left(A_2|A_1\right)+\mathrm{P}\left(\overline{A_1}\right)\mathrm{P}\left(A_2|\overline{A_1}\right)\\&=\tfrac{1}{11}\cdot\tfrac{23}{223}+\tfrac{10}{11}\cdot\tfrac{20}{222}=\quad>\tfrac{1}{11}=\mathrm{P}\left(A_1\right)\end{aligned}$$

דוגמה: בכד 20 כדורים לבנים ו–30 שחורים. מוציאים שני כדורים בזה אחר זה וללא החזרה. מהי ההסתברות למאורע שהראשון לבן והשני שחור?
נסמן ב–A את המאורע שהראשון לבן וב–B את המאורע שהשני שחור. בהתאם לסימון זה, אנו מחפשים את ההסתברות $\mathrm{P}\left(A\cap B\right)$. לפי השיטה הקומבינטורית נקבל:

$$\mathrm{P}\left(A\cap B\right)=\frac{20\cdot30}{50\cdot49}$$

לפי הגישה של הסתברות מותנה נקבל:

$$\begin{gathered}\mathrm{P}\left(A\right)=\tfrac{20}{50}\ ,\ \mathrm{P}\left(B/A\right)=\tfrac{30}{49}\\\Downarrow\\P\left(A\cap B\right)=\tfrac{20}{50}\cdot\tfrac{30}{49}\end{gathered}$$

טענה:

$$\mathrm{P}\left(A\cap B\cap C\right)=\mathrm{P}\left(A\right)\mathrm{P}\left(B|A\right)\mathrm{P}\left(C|A\cap B\right)$$

הוכחה:

$$\mathrm{P}\left(A\cap B\cap C\right)=\mathrm{P}\left(C|A\cap B\right)\mathrm{P}\left(A\cap B\right)=\mathrm{P}\left(C|A\cap B\right)\mathrm{P}\left(A\right)\mathrm{P}\left(B|A\right)$$

■

נמשיך את הדיון בדוגמה האחרונה, ונניח שמוציאים 4 כדורים בזה אחר זה ללא החזרה. נחשב את ההסתברות למאורע שהראשון והשלישי לבנים, והשני והרביעי שחורים. בדרך הקומבינטורית נקבל:

$$\frac{\frac{20!}{18!}\cdot\frac{30!}{28!}}{\frac{50!}{46!}}$$

באמצעות נוסחת ההסתברות נקבל:

$$\mathrm{P}\left(A\cap B\cap C\cap D\right)=\mathrm{P}\left(A\right)P\left(B|A\right)\mathrm{P}\left(C|A\cap B\right)\mathrm{P}\left(D|A\cap B\cap C\right)=\frac{20}{50}\cdot\frac{30}{49}\cdot\frac{19}{48}\cdot\frac{29}{47}$$

דוגמה: נתונות שלוש מגירות. במגירה a שני כדורי זהב, במגירה b כדור זהב וכדור כסף, ובמגירה c שני כדורי כסף. בחרנו מגירה באופן מקרי, ואז בחרנו כדור מקרי מאותה מגירה, נתון כי הוצאנו כדור כסף. נבדוק את ההסתברות לכך שהכדור השני במגירה הוא כדור זהב. כלומר את ההסתברות לכך שהמגירה שנבחרה היא מגירה b. יתכן ובמחשבה ראשונה תעלה האפשרויות כי בהסתברות $\frac{1}{2}$ אך אין זה כך. נשתמש בנוסחת ההסתברות השלמה:

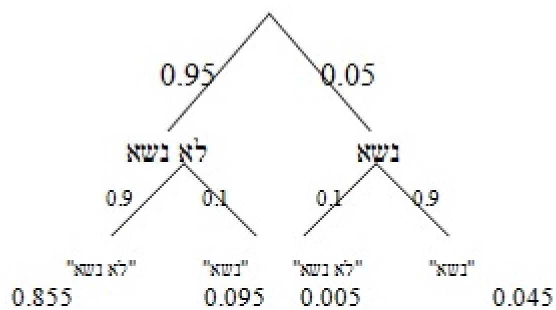
$$\mathrm{P}\left(b|\mathrm{silver}\right)=\frac{\mathrm{P}\left(b\cap\mathrm{silver}\right)}{\mathrm{P}\left(\mathrm{silver}\right)}$$

$$=\frac{\mathrm{P}(\mathrm{silver}|\mathrm{b})\mathrm{P}(b)}{\mathrm{P}(\mathrm{silver}|a)\mathrm{P}(a)+\mathrm{P}(\mathrm{silver}|b)\mathrm{P}(b)+\mathrm{P}(\mathrm{silver}|c)\mathrm{P}(c)}=\frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{0}{2}\cdot\frac{1}{3}+\frac{1}{2}\cdot\frac{1}{3}+\frac{2}{2}\cdot\frac{1}{3}}=\frac{1}{3}$$

ניתן לפתור שאלה זו גם באמצעות עץ.

דוגמה: נתונה אוכלוסייה בה 5% מהפרטים נשאים של נגיף. זהו האפריור, ז`א ההסתברות שאדם מקרי הוא נשא כשלא נתונה שום אנפורמציה נוספת. בהינתן פרט נשא, בדיקת מעבדה קובעת שהוא אכן נשא בהסתברות של 0.9, וקובעת בטעות שהוא לא נשא בהסתברות המשלימה 0.1. בהינתן פרט שאינו נשא, בדיקת המעבדה קובעת שהוא לא נשא בהסתברות של 0.9, וקובעת בטעות שהוא נשא בהסתברות המשלימה 0.1.

בהינתן שבדיקת המעבדה קבעה שפרט כלשהו הוא נשא, מהי ההסתברות שהוא אכן נשא? ייתכן ותשובה פזיזה הייתה קובעת כי ההסתברות היא 0.9, אך אין זה כך. נראה זאת באמצעות עץ (ללא מרכאות מצוין המצב האמיתי. במרכאות מצוין מה נמסר ע`י המעבדה לאחר הבדיקה):

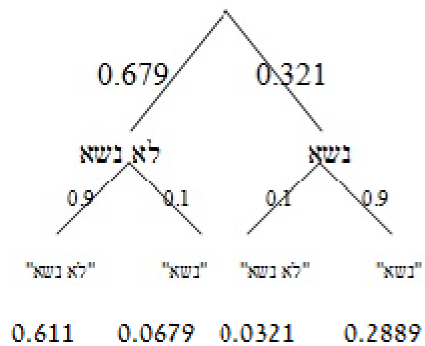


מכאן שההסתברות למאורע שפרט הוא נשא בהינתן שבדיקת המעבדה קבעה שהוא נשא, היא:

$$\frac{0.05\cdot0.9}{0.05\cdot0.9+0.95\cdot0.1}=\frac{0.045}{0.045+0.095}=0.321$$

הסבר: לפני בדיקת המעבדה (אפריורית) חשבנו שהפרט שנדגם באופן מקרי הוא נשא בהסתברות של 0.05. לאחר קבלת התשובה שהוא נשא, עדכנו את ההסתברות (אפוסטריורית) להיות 0.321. נשים לב שההסתברות האפוסטריורית היא פונקציה לא רק של תוצאות הבדיקה אלא גם של ההסתברות האפריורית, ולכן התשובה הפזיזה שגויה.

קיבלנו שההסתברות להיות נשא לאור תשובה של בדיקת המעבדה היא רק 0.321. זו הסתברות נמוכה שלא מצדיקה התחלה של טיפול, ולכן ניתן לזמן את החשודים לבדיקה נוספת. נשים לב שנדגם מקרי מתוך הקבוצה שזומנה שוב הוא בעל הסתברות של 0.321 להיות נשא, ולכן העץ ייראה כך:



ולכן ההסתברות שפרט הוא נשא לאחר שנבדק פעמיים ובשתי הפעמים קיבלת תשובה חיובית היא:

$$\frac{0.321 \cdot 0.9}{0.321 \cdot 0.9 + 0.679 \cdot 0.1} = 0.81$$

באופן כללי, ההסתברות שפרט כלשהו הוא נשא, בהינתן `נשא`, היא:

$$\frac{p \cdot 0.9}{p \cdot 0.9 + (1-p) \cdot 0.1}$$

כאשר p מסמן את ההסתברות שלו להיות נשא טרם ביצוע הבדיקה הנוכחית. בהתאם לזאת נוכל לסמן את ההסתברות שפרט כלשהו הוא נשא לאחר n בדיקות ב-p_n, ואז:

$$p_{n+1} = \frac{p_n \cdot 0.9}{p_n \cdot 0.9 + (1-p_n) \cdot 0.1}$$

כאשר $p_0 = 0.05$.
ניתן לראות שמדובר בסדרה מונוטונית עולה ב-n, המתכנסת ל-1 (ולא למשל ל-0.9), באמצעות העובדה ש-1 הוא הפתרון החיובי היחיד למשוואה:

$$x = \frac{x \cdot 0.9}{x \cdot 0.9 + (1-x) \cdot 0.1}$$

כלומר לכל הסתברות $p < 1$ שנבחר קיים n מספיק גדול (דהיינו מספר בדיקות מספיק גדול) כך שאם נבדק נמצא n פעמים רצופות `נשא`, אז ההסתברות שהוא אכן נשא היא לפחות p.

דוגמה: נתונות שלוש כוסות, ובאחת מהן בלבד מניחים מטבע. מהמר מנחש באיזו כוס המטבע נמצא, ולפני שמספרים לו על תוצאות הניחוש שלו, חושפים בפניו כוס ריקה

אחת מבין השתיים הנותרות. שים לב כי הדבר תמיד ניתן. בפרט, אם יש שתי כוסות ריקות, הבחירה ביניהן היא אקראית. כעת המהמר יודע שהמטבע נמצא באחת משתי כוסות – זו שבחר מלכתחילה או זו שלא נחשפה בפניו, והוא מקבל אפשרות להמר מחדש על אחת משתי הכוסות הללו. האם כדאי למהמר לשנות את ההימור במקורו או לדבוק בו? על פניו נראה כי אין זה משנה וכל מה שיעשה הסתברות הצלחתו היא 0.5 אך לא כן הוא.

נשים לב שמי שלא משנה את הימורו זוכה במטבע בהסתברות שליש, כי הוא זוכה אך ורק אם בחר מלכתחילה את הכוס הנכונה, וההסתברות לכך היא $\frac{1}{3}$. לעומת זאת מי שמשנה את הימורו זוכה בהסתברות שני שליש, כי הוא זוכה אך ורק אם בחר מלכתחילה כוס לא־נכונה, וההסתברות לכך היא $\frac{2}{3}$.

13.3 שכיחות יחסית מותנה

שכיחות יחסית יכולה להשתנות כאשר נתון שהתרחש מאורע כלשהו. למשל, נניח שכך מתפלגות יחידות הלימוד במתמטיקה, בבית־ספר של 60 בנים ו–40 בנות:

	3 יחידות	4 יחידות	5 יחידות	סה`כ
בנים	10	30	20	60
בנות	10	20	10	40
סה`כ	20	50	30	

נשים לב שהשכיחות היחסית של הבנים משתנה, אם למשל נתון שהפרט נלקח מקבוצה מסוימת:

$$f\left(\text{boy}\right) = \frac{60}{100} = 0.6$$

$$f\left(\text{boy}|3\,\text{units}\right) = \frac{10}{20} = 0.5$$

14 מאורעות בלתי תלויים

הגדרה: המאורע A נקרא **בלתי־תלוי** במאורע B אם מתקיים:

$$\boxed{\mathrm{P}\left(A|B\right) = \mathrm{P}\left(A\right)}$$

טענה: אם A בלתי־תלוי ב–B אז גם B בלתי־תלוי ב–A.

הוכחה: נתון ש–A בלתי־תלוי ב–B, ולכן:

$$\mathrm{P}\left(A\right) = \mathrm{P}\left(A|B\right) = \frac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)}$$
$$\Updownarrow$$
$$\mathrm{P}\left(A\right)\cdot \mathrm{P}\left(B\right) = \mathrm{P}\left(A\cap B\right)$$

נשים לב שפעולת כפל חילופית, וכן חיתוך מקיים $A\cap B=B\cap A$, ולכן ניתן להסיק שגם B בלתי־תלוי ב־A. ■

הגדרה כללית: נגדיר באופן כללי שהמאורעות A,B בלתי־תלויים אם מתקיים:

$$\boxed{\mathrm{P}\left(A\right)\cdot\mathrm{P}\left(B\right)=\mathrm{P}\left(A\cap B\right)}$$

הגדרה שקולה: נשים לב שההגדרה לאי־תלות שקולה להגדרה הבאה:

$$\boxed{\mathrm{P}\left(A|B\right)=\mathrm{P}\left(A|\overline{B}\right)}$$

הוכחה: לפי נוסחת ההסתברות השלמה מתקיים:

$$\mathrm{P}\left(A\right)=\mathrm{P}\left(B\right)P\left(A|B\right)+\mathrm{P}\left(\overline{B}\right)\mathrm{P}\left(A|\overline{B}\right)$$

ולכן אם מתקיים השוויון הנ`ל נוכל להסיק שהוא שקול להגדרה של אי־תלות:

$$\mathrm{P}\left(A\right)=P\left(A|B\right)\left[\mathrm{P}\left(B\right)+\mathrm{P}\left(\overline{B}\right)\right]=\mathrm{P}\left(A|B\right)$$

■

הערה: ניתן לראות שכאשר A,B בלתי־תלויים אז גם המשלימים $\overline{A},\overline{B}$ בלתי־תלויים:

$$\begin{gathered}
\mathrm{P}\left(A\cap B\right)=\mathrm{P}\left(A\right)\mathrm{P}\left(B\right)\\
\Downarrow\\
1-\mathrm{P}\left(\overline{A\cap B}\right)=\left(1-\mathrm{P}\left(\overline{A}\right)\right)\left(1-\mathrm{P}\left(\overline{B}\right)\right)\\
\Downarrow\\
\mathrm{P}\left(\overline{A\cap B}\right)=\mathrm{P}\left(\overline{A}\right)+\mathrm{P}\left(\overline{B}\right)-\mathrm{P}\left(\overline{A}\right)\mathrm{P}\left(\overline{B}\right)\\
\Downarrow\\
\mathrm{P}\left(\overline{A}\right)+\mathrm{P}\left(\overline{B}\right)-\mathrm{P}\left(\overline{A\cap B}\right)=\mathrm{P}\left(\overline{A}\right)\mathrm{P}\left(\overline{B}\right)\\
\Downarrow\\
\mathrm{P}\left(\overline{A}\right)+\mathrm{P}\left(\overline{B}\right)-\mathrm{P}\left(\overline{A}\cup\overline{B}\right)=\mathrm{P}\left(\overline{A}\right)\mathrm{P}\left(\overline{B}\right)\\
\Downarrow\\
\mathrm{P}\left(\overline{A}\cap\overline{B}\right)=\mathrm{P}\left(\overline{A}\right)\mathrm{P}\left(\overline{B}\right)
\end{gathered}$$

המעבר הרביעי נובע מכללי דה־מורגן, והמעבר החמישי נובע מהנוסחה:

$$\mathrm{P}\left(A\cup B\right)=\mathrm{P}\left(A\right)+\mathrm{P}\left(B\right)-\mathrm{P}\left(A\cap B\right)$$

סיכום: התנאים הבאים שקולים וכולם מגדירים אי־תלות בין שני מאורעות A,B:

1. $\mathrm{P}\left(A|B\right)=\mathrm{P}\left(A\right)$

2. $\mathrm{P}(B|A) = \mathrm{P}(B)$
3. $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$
4. $\mathrm{P}(A|B) = \mathrm{P}\left(A|\overline{B}\right)$
5. $\mathrm{P}\left(\overline{A} \cap \overline{B}\right) = \mathrm{P}\left(\overline{A}\right)\mathrm{P}\left(\overline{B}\right)$

דוגמה: זורקים שתי קוביות הוגנות. נגדיר את המאורעות:

$$A = \text{"even number on the first"}$$

$$B = \text{"even number on the second"}$$

נשים לב שמתקיים:

$$\mathrm{P}(A) = \mathrm{P}(B) = 0.5$$

$$\mathrm{P}(A \cap B) = \tfrac{1}{4}$$

ולכן נסיק:

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2} = \mathrm{P}(A)$$

מכאן שהמאורעות A, B בלתי־תלויים.

נגדיר מאורעות נוספים:

$$C = \text{"3 on the first"}$$

$$D = \text{"the sum is 7"}$$

נשים לב שמתקיים:

$$\mathrm{P}(C) = \mathrm{P}(D) = \tfrac{1}{6}$$

$$\mathrm{P}(C \cap D) = \tfrac{1}{36}$$

ולכן נסיק:

$$\mathrm{P}(C|D) = \frac{\mathrm{P}(C \cap D)}{\mathrm{P}(D)} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6} = \mathrm{P}(C)$$

מכאן שהמאורעות C, D בלתי־תלויים.

נגדיר מאורע נוסף:

$$E = \text{"the sum is 9"}$$

נשים לב שמתקיים:

$$\mathrm{P}(E) = \tfrac{1}{9}$$

$$\mathrm{P}(C|E) = \tfrac{1}{4} > \tfrac{1}{6}$$

$$\mathrm{P}(E|C) = \tfrac{1}{6} > \tfrac{1}{9}$$

נשים לב שיש כאן טענה כללית: אם $\mathrm{P}(A|B) > \mathrm{P}(A)$, אז $\mathrm{P}(B|A) > \mathrm{P}(B)$. במילים: אם B מגדיל (מקטין) את ההסתברות של A, אז גם A מגדיל (מקטין) את ההסתברות של B.

הוכחה:

$$\mathrm{P}(A) < \mathrm{P}(A|B) = \tfrac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)} = \tfrac{\mathrm{P}(A)\mathrm{P}(B|A)}{\mathrm{P}(B)}$$
$$\Updownarrow$$
$$1 < \tfrac{\mathrm{P}(B|A)}{\mathrm{P}(B)}$$
$$\Updownarrow$$
$$\mathrm{P}(B) < \mathrm{P}(B/A)$$

הערה: נזכור שלפי נוסחת ההסתברות השלמה מתקיים:

$$\mathrm{P}(A) = \mathrm{P}(B)\,\mathrm{P}(A|B) + \mathrm{P}\left(\overline{B}\right)\mathrm{P}\left(A|\overline{B}\right)$$

כלומר מבצעים מיצוע משוקלל של ההסתברויות המותנות, לפי החלוקה של מרחב המדגם ל-B ו-$\overline{B}$. במקרה שבו A, B בלתי־תלויים המשקוללים שווים, כי $\mathrm{P}(A|B) = \mathrm{P}(A|B) = \mathrm{P}(A)$.

דוגמה: זורקים שתי קוביות הוגנות. נגדיר את המאורעות:

$$A = \text{"even number on the first"}$$

$$B = \text{"even number on the second"}$$

$$C = \text{"the sum is even"}$$

נשים לב שמתקיים:

$$\mathrm{P}(A) = \mathrm{P}(B) = \mathrm{P}(C) = \tfrac{1}{2}$$

$$\mathrm{P}(A \cap B) = \tfrac{1}{4}$$

כמו־כן הראינו לעיל ש-A, B בלתי־תלויים, וכן גם A, C בלתי־תלויים וגם B, C. לעומת זאת מתקיים:

$$\frac{1}{4} = \mathrm{P}(A \cap B \cap C) = \mathrm{P}(A \cap B) \neq \mathrm{P}(A)\,\mathrm{P}(B)\,\mathrm{P}(C) = \frac{1}{8}$$

כלומר A, B, C בלתי־תלויים בזוגות, אך לא בלתי־תלויים בשלשות.

הגדרה: המאורעות $\{A_i\}_{i=1}^{m}$ הם בלתי־תלויים, אם לכל קבוצה חלקית של מאורעות מגודל k, $1 \le k \le n$, $A_{i_1}, A_{i_2}, ..., A_{i_k}$, מתקיים:

$$\mathrm{P}\left(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_k}\right) = \mathrm{P}\left(A_{i_1}\right) \cdot \mathrm{P}\left(A_{i_2}\right) \cdot ... \cdot \mathrm{P}\left(A_{i_k}\right)$$

כך למשל שלושת המאורעות A, B, C בלתי־תלויים אם מתקיימים <u>כל</u> ארבעת התנאים הבאים:

1.
$$\mathrm{P}\left(A \cap B \cap C\right) = \mathrm{P}\left(A\right) \mathrm{P}\left(B\right) \mathrm{P}\left(C\right)$$

2.
$$\mathrm{P}\left(A \cap B\right) = \mathrm{P}\left(A\right) \mathrm{P}\left(B\right)$$

3.
$$\mathrm{P}\left(A \cap C\right) = \mathrm{P}\left(A\right) \mathrm{P}\left(C\right)$$

4.
$$\mathrm{P}\left(B \cap C\right) = \mathrm{P}\left(B\right) \mathrm{P}\left(C\right)$$

הגדרה שקולה: המאורעות $\{A_i\}_{i=1}^{m}$ הם בלתי־תלויים, אם לכל $I, J \subset \{1, ..., n\}$, $I \cap J = \phi$ מתקיים:

$$\mathrm{P}\left(\underset{i\in I}{\cap} A_i | \underset{i\in J}{\cap} A_i\right) = \mathrm{P}\left(\underset{i\in I}{\cap} A_i\right)$$

דוגמה: נניח שכך מתפלגות יחידות הלימוד במתמטיקה, בבית־ספר של 60 בנים ו–40 בנות:

	3 יחידות	4 יחידות	5 יחידות
בנים	10	20	10
בנות	15	30	15

נשים לב שבדוגמה זו מתקיים:

$$\mathrm{P}\left(4\,\mathrm{units}\right)=\frac{50}{100}=\frac{1}{2}$$

$$\mathrm{P}\left(4\,\mathrm{units}|\mathrm{boys}\right)=\frac{20}{40}=\frac{1}{2}$$

ולכן המאורעות `4 יחידות` ו`בנים` הם בלתי־תלויים. לכאורה דוגמה זו תומכת בהצעה להגדיר אי־תלות בין מאורעות בהקשר של שכיחות יחסית באופן דומה להגדרתה בהקשר של הסתברות. אך זה לא כך.
נתבונן בהגדרה זו ביחס לשכיחות מותנה:

$$\mathrm{P}\left(A|B\right)=\frac{\mathrm{P}\left(A\cap B\right)}{\mathrm{P}\left(B\right)}\overset{?}{\gtreqless}P\left(A\right)$$

כדי לקבל שוויון צריכות לקרות התרחשויות נדירות. למשל, נניח שנדגמו 101 ילדים, ונרשמו המין והאם הם אוהבים לשתות שוקו. התוצאות שהתקבלו הן:

	drink cacao	not drink cacao	
boys	x	$61-x$	61
girls	$51-x$	$x-11$	40
	51	50	101

כדי לקבל מצב של אי־תלות בין שתיית שוקו למין בהתאם להגדרה בהקשר של הסתברות, נצטרך שהשכיחות היחסית של שותי השוקו בקרב הבנים תהיה שווה לשכיחות היחסית של שותות השוקו בקרב הבנות, ולכן בעצם שווה לשכיחות היחסית של שותי השוקו באוכלוסייה הכללית. נבדוק איזה x מקיים את הדרישה, בהתאם לסימון של x כמספר שותי השוקו בקרב הבנים:

$$\frac{x}{61}=\frac{51}{100}$$
$$\Downarrow$$
$$x=\frac{51}{100}\cdot 61=31.11$$

כמובן לא ייתכן ש־ 31.11 בנים שותים שוקו, ולכן אין מצב שבו המאורעות הללו בלתי־תלויים.

14.0.1 דוגמה: אוניברסיטת ברקלי

נתונים לגבי מגישי מועמדות ללימודים וקבלה לפי מין, בשנת 1973:

	men	woman	
received	1198	557	1755
not received	1493	1278	2771
	2691	1835	4526

מהנתונים הללו נובע ש־44% מהגברים התקבלו בעוד שרק 30% מהנשים התקבלו.
נתונים אלה עוררו טענה שקיימת אפליה כנגד נשים באוניברסיטת ברקלי.

באופן מפתיע, כאשר בוצע פילוח נוסף של הנתונים לפי חוגי לימוד שאליהם נרשמו המועמדים/ות, התקבלו התוצאות הבאות:

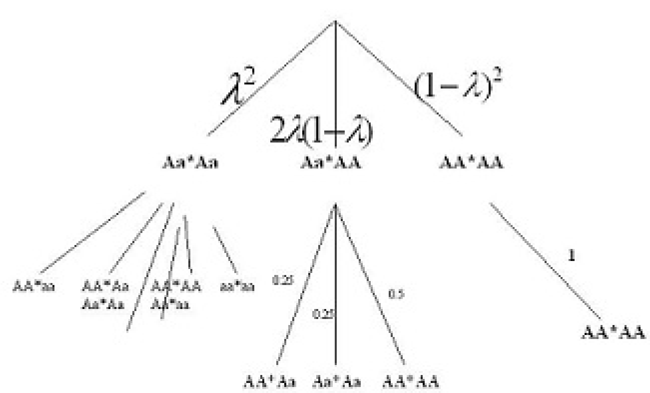
department	candidates - men	received - men(%)	candidates - women	received - women (%)
A	825	62%	108	82%
B	560	63%	25	68%
C	325	37%	593	34%
D	417	33%	375	35%
E	191	28%	393	24%
F	373	6%	341	7%

כאשר בוחנים כל מחלקה לחוד, ניתן לראות כי בחלק מהמחלקות דווקא אחוז הנשים המתקבלות גבוה מהערך המקביל אצל הגברים. פרדוקס? מתגלה שנשים הגישו יותר מועמדויות לחוגים שבהם אחוזי הקבלה מראש היו נמוכים, ביחס לגברים שהגישו יותר מועמדויות לחוגים שבהם אחוזי הקבלה גבוהים. לכן, הסיבה לפער בין אחוזי הקבלה של נשים וגברים אינו נובע מאפליה על רקע מגדרי, אלא מאופי החוג שאליו הגישו מועמדות הנשים והגברים. נציין רק שתהליך הקבלה הוא לכל מחלקה בנפרד ואין, למשל, צורך להתקבל ראשית לאוניברסיטה.

14.0.2 דוגמה: גנטיקה

נניח שהגן a הוא גן קטלני. באוכלוסייה מסוימת כל פרט מאופיין באמצעות אחד מהמצבים הבאים AA (נקי), Aa (נשא), או aa (מת ממחלה). פרופורציית הנשאים Aa מבין הפרטים החיים היא λ, כך שפרופורציית הנקיים AA היא $1-\lambda$. בהסתברות $(1-\lambda)^2$ נקבל שני הורים AA, כך שאז גם הצאצא יהיה AA (בהסתברות 1). בהסתברות λ^2 נקבל שני הורים Aa, כך שהצאצא יהיה Aa בהסתברות 0.5 או AA בהסתברות 0.25 או aa בהסתברות 0.25. כלומר ההסתברות שתינוק אקראי יהיה חולה וימות (aa) היא $\frac{1}{4}\lambda^2$. בהסתברות $2\lambda(1-\lambda)$ (הכפלנו ב-2 כי הסדר משנה) נקבל הורה אחד AA והורה אחד Aa, כך שהצאצא יהיה AA בהסתברות 0.5 או Aa בהסתברות 0.5.

נבדוק מהי ההסתברות לנשאות כאשר מכליאים בין קרובי־משפחה. למשל אח ואחות, אך עדיין נניח שהורי שני בני הזוג הם מקריים.
נציג את ההסתברויות באמצעות עץ:



החלק העליון של העץ מייצג את ההסתברויות בקרב ההורים של בני הזוג, כפי שחישבנו לעיל. ההסתעפויות בהמשך מייצגות את האח והאחות. נגדיר את המאורעות הבאים עבור האח והאחות:

$$A = \text{"both carriers"}$$

$$\overline{B} = \text{"at least one of them died"}$$

נחשב את המאורע ששניהם נשאים, בהינתן ששניהם לא מתו (אחרת לא יכלו להביא צאצאים בהמשך):

$$\mathrm{P}\left(A|B\right) = \frac{\mathrm{P}\left(A\cap B\right)}{\mathrm{P}\left(B\right)}$$

$$= \frac{\frac{1}{4}\lambda^2+\frac{1}{4}\cdot 2\lambda(1-\lambda)}{\frac{1}{4}\lambda^2+\frac{1}{4}\cdot 2\lambda(1-\lambda)+\frac{1}{16}\lambda^2+\frac{1}{4}\lambda^2+\frac{1}{2}\lambda(1-\lambda)+\lambda(1-\lambda)+(1-\lambda^2)}$$

$$\cong \frac{\frac{1}{2}\lambda}{1-\frac{7}{16}\lambda}$$

שים לב כי בקירוב האחרון החלפנו את λ^2 באפס. זאת אומרת, עבור λ קטן (שזה המקרה הטיפוסי) מתקיים הקירוב $\mathrm{P}\left(A|B\right) \cong \frac{1}{2}\lambda$. כלומר, ההסתברות לזוג נשאים עולה מ-λ^2 מזיווג מקרי, ל-$\frac{\lambda}{2}$ בקירוב מזיווג בין אח ואחות.

14.1 אי תלות מותנה

הגדרה: המאורעות A ו־ B יקראו בלתי תלויים בהנתן C אם מתקיים

$$\mathrm{P}(A\cap B|C) = \mathrm{P}(A|C)\cdot \mathrm{P}(B|C)$$

במילים, אם התרחש המאורע C אזי המאורעות A ו־ B הינם בלתי תלויים.

כמובן, אין זה אומר ש A ו־ B בלתי תלויים. יתר על כן, אין זה גורר כי A ו־ B בלתי תלויים בהינתן $\overline{C}$.

דוגמה: נחזור לדוגמה בעמוד 74 , ראינו שם כי תוצאה חיובית עדיין רחוקה מלשכנע כי הנבדק הוא אכן נשא. מקובל לזמן נבדק שכזה לבדיקה נוספת. נסמן ב־C את המאורע נשא, ב־A את המאורע נשא בבדיקה הראשונה ו־B נשא בבדיקה שניה. כמובן שתוצאות הבדיקות אינן בלתי תלויות. אולם ניתן להניח כי בהנתן C (ובמקרה זה אף $\overline{C}$) מאורעות A ו־ B בלתי תלויים.

בפרט: $\mathrm{P}(A\cap B|C) = \mathrm{P}(A|C)\cdot \mathrm{P}(B|C) = 0.9\cdot 0.9 = 0.81$

נשאל כעת את השאלה מהו $\mathrm{P}(C|A\cap B)$

$$\mathrm{P}(C|A\cap B) = \frac{\mathrm{P}(A\cap B\cap C)}{\mathrm{P}(A\cap B)} = \frac{\mathrm{P}(A\cap B|C)\cdot \mathrm{P}(C)}{\mathrm{P}(A\cap B|C)\cdot \mathrm{P}(C)+\mathrm{P}(A\cap B|\overline{C})\cdot \mathrm{P}(\overline{C})} =$$

$$\frac{\mathrm{P}(A|C)\cdot \mathrm{P}(B|C)\cdot \mathrm{P}(C)}{\mathrm{P}(A|C)\cdot \mathrm{P}(B|C)\cdot \mathrm{P}(C)+\mathrm{P}(A|\overline{C})\cdot \mathrm{P}(B|\overline{C})\cdot \mathrm{P}(\overline{C})} = \frac{0.9\cdot 0.9\cdot 0.5}{0.9\cdot 0.9\cdot 0.05+0.1\cdot 01\cdot 0.95} = 0.81$$

בפרט אנו רואים כי לאחר שתי בדיקות עוקבות שתוצאותיהן חיוביות ההסתברות כי הנבדק אכן נשא היא גבוהה.

הערה: ניתן לראות כי אם A ו־ B בלתי תלויים בהנתן C אזי: $\mathrm{P}(A|B\cap C)=\mathrm{P}(A|C)$ ולמעשה זוהי הגדרה שקולה לאי תלות מותנה (הוכח)

15 שאלות חזרה

1. על פי הגדרה של הסתברות מותנה קיבלנו $\mathrm{P}(A\cap B)=P(A/B)P(B)$, הוכיחו באינדוקציה טענה כללית יותר, טענה זו נקראת גם `כלל ההכפלה` או `כלל השרשרת`:

   לכל אוסף מאורעות $A_1....A_n$ מתקיים:

$$\mathrm{P}(\cap_{i=1}^{n}A_i)=\mathrm{P}(A_1)\cdot\mathrm{P}(A_2|A_1)\cdot\mathrm{P}(A_3|A_2\cap A_1)....\mathrm{P}(A_n|A_{n-1}\cap A_{n-2.....}\cap A_1)$$

2. שישה אנשים משחקים רולטה רוסית. מעבירים אקדח עם כדור אחד אמיתי וחמישה סרק, מקומו של הכדור אינו ידוע. כל אחד בתורו לוחץ על ההדק. האם עדיף לשחק ראשון או אחרון? (הוכיחו את תשובתכם באופן פורמלי, העזרו בכלל השרשרת מהשאלה הקודמת)

3. יהי מרחב הסתברות עם מרחב מדגם Ω ופונקצית הסתברות P. הוכח שעבור כל זוג מאורעות $A,B\subseteq\Omega$ בעלי הסתברות חיובית מתקיים: $\mathrm{P}\left(A\cap B|A\right)\geq\mathrm{P}\left(A\cap B|A\cup B\right)$

4. יהי מרחב הסתברות עם מרחב מדגם Ω ופונקצית הסתברות P. הוכח או הפרך (כלומר, תנו דוגמא נגדית): לכל שלשה של מאורעות $A,B,C\subseteq\Omega$ בעלי הסתברות חיובית מתקיים $\mathrm{P}(A\cap B|A)\geq\mathrm{P}(A\cap B|A\cap C)$

5. יהי מרחב הסתברות עם מרחב מדגם Ω ופונקצית הסתברות P. הוכח או הפרך: לכל שלשה של מאורעות A B ו־ C מתקיים: $\mathrm{P}(A|B\cap C)=\frac{\mathrm{P}(A\cap B|C)}{\mathrm{P}(B|C)}$

6. יהי מרחב הסתברות עם מרחב מדגם Ω ופונקצית הסתברות P. הוכח או הפרך: לכל זוג מאורעות A ו־ B מתקיים: $\mathrm{P}(A|B)=\mathrm{P}(B|A)$

7. באי מסוים שלושה שבטים: שבט דוברי אמת המהווה 40% מאוכלוסיית האי, שבט דוברי שקר המהווה 20% מאוכלוסיית האי ושבט שלישי אשר אנשיו אומרים את האמת בהסתברות 0.6. מה ההסתברות כי אם אשאל תושב מקרי שאלה אחת אקבל תשובה שקרית?

8. אבי ובתיה מוצאים שק עם $2n+1$ מטבעות זהב. בתיה מציעה את המשחק הבא על מנת לקבוע מי יקבל את כל המטבעות. בתיה תטיל $n+1$ מטבעות ואבי את ה־n הנותרים. לאחר מכן הם יספרו את מספרות תוצאות ה`עץ` שכל אחד קיבל, ואם בתיה קיבלה יותר `עץ` אז היא זוכה בכל המטבעות, אחרת אבי זוכה בכל המטבעות. אבי טוען כי המשחק אינו הוגן ואילו בתיה טוענת כי הוא כן, מי צודק?

9. לרשותנו מטבע הוגן ושני כדים עם כדורים:

   - כד א' מכיל 3 כדורים אדומים ו־2 כחולים

- כד ב' מכיל 2 כדורים אדומים ו־8 כחולים

מטילים את המטבע ואם יצא `עץ` מוציאים כדור מכד א', ואם יצא `פלי` מוציאים כדור מכד ב'. ענו על השאלות הבאות

(א) בהנתן שיצא `עץ`, מה ההסתברות שיצא כדור אדום?

(ב) בהנתן שיצא `פלי` מה ההסתברות שיצא כדור אדום?

(ג) מה ההסתברות שיצא כדור אדום?

(ד) מה ההסתברות שיצא כדור כחול?

(ה) בהנתן שיצא כדור כדור אדום, מה ההסתברות שיצא עץ?

(ו) בהנתן שיצא כדור כדור כחול, מה ההסתברות שיצא עץ?

הערה: לפתרון השאלה הגדירו את המאורעות המתאימים ורשמו באופן מלא את התכונות בהן אתם משתמשים

10. עבור זוג מאורעות A ו־B ידוע כי $\mathrm{P}(A)=0.3$ ו־$\mathrm{P}(A\cup B)=0.8$.

(א) הוכיחו כי $\mathrm{P}(B)\geq 0.5$.

(ב) חשבו את $\mathrm{P}(B)$ אם ידוע ש A ו־B זרים.

(ג) חשבו את $\mathrm{P}(B)$ אם ידוע ש A ו־B ב`ת.

11. למשפחה יש $n\geq 2$ ילדים. נגדיר שני מאורעות:

- A ־ למשפחה יש ילדים משני המינים.
- B ־ למשפחה יש לכל היותר בת אחת.

נניח כי ההסתתברות ללדת בן או בת היא שווה, וכי הלידות הן ב`ת.

איזה אחת מהטענות הבאות היא הנכונה:

(א) המאורעות A ו־B הינם ב`ת לכל n.

(ב) המאורעות A ו־B הינם תלויים לכל n.

(ג) קיים ערך יחיד של n עבורו A ו־B הינם ב`ת.

(ד) אף אחת מהתשובות א־ג אינה נכונה.

הדרכה: חשבו את $\mathrm{P}(A)$, $\mathrm{P}(B)$ ו־$\mathrm{P}(A\cap B)$ עבור ערכים קטנים של n ובדקו את הטענות. לאחר מכן נסו למצוא נוסחא כללית להסתברויות הנ`ל וענו בעזרתה על השאלה.

12. עבור כל אחת מהטענות הבאות ציינו האם היא נכונה או לא. במקרה שכן הוכיחו את הטענה, ובמקרה שלא הציגו דוגמא נגדית להפרכתה:

(א) אם A ו־B ב`ת, אז $A$ ו־$\overline{B}$ ב`ת.

(ב) אם A ו־B ב`ת ו־$A$ ו־$C$ ב`ת, אז A ו־$(B\cup C)$ ב`ת.

(ג) אם A ו־B ב`ת ו־$A$ ו־$C$ ב`ת, אז B ו־C ב`ת.

(ד) אם A ו־$(B\cap C)$ ב`ת, אז $A$ ו־$\left(\overline{B}\cup\overline{C}\right)$ ב`ת.

(ה) אם A בלתי תלוי ב־B, ו־B בלתי תלוי ב־C, אז גם A בלתי תלוי ב־C

חלק V

משתנים מקריים

16 מפונקציית הסתברות לפונקציית התפלגות

הגדרה: נתון מרחב מדגם כלשהו Ω. נאמר ש־X הוא **משתנה מקרי** (מ`מ) אם לכל מאורע פשוט $\omega \in \Omega$ הוא מתאים מספר ממשי. למעשה, X היא פונקציה ממרחב המדגם למספרים ממשיים.

הגדרה: **פונקציית ההתפלגות** (או ההסתברות) של משתנה מקרי X, מגדירה לכל x מספר ממשי אחר שמסומן $\mathrm{P}\left(X=x\right)$. כלומר, זו ההסתברות שהמשתנה המקרי X שווה ל־x מסוים. נשים לב שהתפלגות של משתנה מקרי מגדירה **חלוקה** של מרחב המדגם. כלומר המקורות השונים של כל הערכים האפשריים השונים של משתנה מקרי מכסים את מרחב המדגם כולו והם זרים בזוגות.

דוגמה: הטלת מטבע פעמיים. נסמן:

$$\Omega=\{HH,HT,TH,TT\}$$

$$X\left(HH\right)=2$$

$$X\left(TH\right)=1$$

$$X\left(HT\right)=1$$

$$X\left(TT\right)=0$$

כלומר, המשתנה המקרי X בדוגמה זו סופר את מספר המופעים של `ראש` בהטלת מטבע פעמיים.

המשתנה המקרי X מגדיר מאורעות. כך למשל $X=0$ מגדיר את המאורע $\{TT\}$, $X=1$ מגדיר את המאורע $\{TH,HT\}$ ו־$X=0$ מגדיר את המאורע $\{TT\}$. נשים לב שניתן גם להגדיר $\{X<2\}=\{HT,TH,TT\}$.

בדוגמה זו נקבל את ההסתברויות של המאורעות הבאים:

$$\mathrm{P}\left(X=2\right)=0.25$$

$$\mathrm{P}\left(X=1\right)=0.5$$

$$\mathrm{P}\left(X=0\right)=0.25$$

סכום ההסתברויות הוא בהכרח 1.

דוגמה: במועצה מקומית 5 נציגים. ראובן ושמעון – מהליכוד; לוי, יהודה ויששכר – מהעבודה. יש להרכיב ועדה בת שני חברים, וכל שילוב של שניים מהחמישה הוא

מאורע אפשרי במרחב המדגם. נגדיר $X = \text{number of Likud members in the committee}$,
ובהתאם להגדרה זו:

$$X\left(\text{Reuven and Shimon}\right) = 2$$

$$X\left(\text{Reuven and Levi}\right) = 1$$

$$\vdots$$

נקבל את ההסתברויות הבאות:

$$\mathrm{P}\left(X = 0\right) = \frac{\binom{3}{2}}{\binom{5}{2}} = \frac{3}{10}$$

$$\mathrm{P}\left(X = 1\right) = \frac{\binom{3}{1}\binom{2}{1}}{\binom{5}{2}} = \frac{6}{10}$$

$$\mathrm{P}\left(X = 2\right) = \frac{\binom{2}{2}}{\binom{5}{2}} = \frac{1}{10}$$

16.1 פונקציית התפלגות מצטברת

הגדרה: נאמר ש-$F_X\left(x\right)$ היא **פונקציית התפלגות מצטברת** של המשתנה המקרי X, אם היא מחזירה את ההסתברות ש-X קטן או שווה לערך x. כלומר:

$$F_X\left(x\right) = \mathrm{P}\left(X \leq x\right)$$

פונקציה זו היא מונוטונית לא־יורדת ב-x, כי ערכי הסתברות הם תמיד אי־שליליים.

תכונות $F_X\left(x\right)$:

1. מונוטונית לא־יורדת ב-x
2. רציפה מימין
3.

$$\lim_{x \to \infty} F_X\left(x\right) = 1$$

$$\lim_{x \to -\infty} F_X\left(x\right) = 0$$

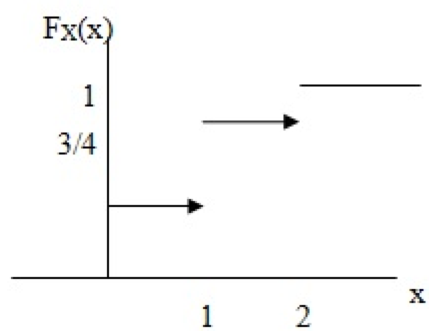
דוגמה: נתבונן בדוגמה שהזכרנו על הטלת מטבע פעמיים. נשרטט את גרף פונקציית ההתפלגות המצטברת של המשתנה המקרי X שמוגדר כמספר הפעמים שמתקבל `ראש`. מתקיים:

$$\mathrm{P}\left(X\leq 0\right)=0.5$$

$$\mathrm{P}\left(X\leq 1\right)=0.25+0.5=0.75$$

$$\mathrm{P}\left(X\leq 2\right)=1$$

ולכן גרף הפונקציה הוא:



17 התפלגויות

17.1 התפלגות ברנולי

(משפחה חד־פרמטרית של התפלגויות)
מ`מ X ייקרא מפולג ברנולי אם הוא מהצורה $X=0$ או $X=1$, כאשר $\mathrm{P}\left(X=0\right)=p$ ו־$\mathrm{P}\left(X=1\right)=1-p$. עבור $0\leq p\leq 1$ כלשהו. מסמנים משתנה מקרי ברנולי $X\sim B\left(p\right)$.

מקרה פרטי של התפלגות זו הוא `אינדיקטור` (או `מציין`) של מאורע כלשהו A:

$$X\left(\omega\right)=\begin{cases}1 & \omega\in A\\ 0 & \omega\in\overline{A}\end{cases}$$

ומכאן $p=\mathrm{P}\left(X=1\right)=\mathrm{P}\left(A\right)$.

17.2 התפלגות אחידה

(משפחה דו־פרמטרית של התפלגויות)
מ`מ X ייקרא מפולג אחיד על קטע $[a,b]$ (עבור $a\leq b$), כאשר a,b מספרים שלמים כלשהם (ייתכן גם שליליים), אם פונקציית ההסתברות שלו היא:

$$\mathrm{P}\left(X=k\right)=\begin{cases}\frac{1}{b-a+1} & k\in[a,b]\\ 0 & \text{otherwise}\end{cases}$$

(עבור k שלם)
מסמנים משתנה מקרי אחיד $X \sim U\left[a,b\right]$.

17.3 התפלגות בינומית

(משפחה דו־פרמטרית של התפלגויות)
נניח כי n מספר טבעי וכי p שבר כלשהי בקטע $[0,1]$.
מ`מ X ייקרא מפולג בינומית עם פרמטרים n,p, אם עבור כל k שלם, $0 \le k \le n$, פונקציית ההסתברות שלו היא:

$$\mathrm{P}\left(X=k\right)=\left(\begin{array}{c} n\\ k \end{array}\right)p^{k}\left(1-p\right)^{n-k}$$

מסמנים משתנה מקרי בינומי $X \sim Bin\left(n,p\right)$.

נוכיח שפונקציה זו אכן מגדירה התפלגות:

ראשית ניכר שטווח ערכי הפונקציה הוא מספרים אי־שליליים.
שנית, לפי הבינום של ניוטון מתקיים:

$$\sum_{k=1}^{n}\mathrm{P}\left(X=k\right)=\sum_{k=0}^{n}\left(\begin{array}{c} n\\ k \end{array}\right)p^{k}\left(1-p\right)^{n-k}=\left(p+\left(1-p\right)\right)^{n}=1$$

הסבר להתפלגות הבינומית:

נניח שמנקודת מבטו של מי שבוחן ניסוי מסוים יש שתי תוצאות אפשריות (למרות שיתכן ולניסוי עצמו יש מספר רב של אפשרויות). נקרא לתוצאה אפשרית אחת הצלחה ולאחרת כישלון. נניח כי ההסתברות להצלחה היא p, ומכאן שההסתברות המשלימה לכישלון היא $1-p$. חוזרים על הניסוי n פעמים באופן בלתי־תלוי. נשים לב שלמשל כי אם $n=3$ נקבל את ההסתברויות:

$$\mathrm{P}\left(\text{failure,success,success}\right)=p^{2}\left(1-p\right)$$

$$\mathrm{P}\left(\text{success,success,failure}\right)=p^{2}\left(1-p\right)$$

במאורעות אלו יש שתי הצלחות וכישלון אחד. כך גם בכל המקרים האחרים בהם נבחן את ההסתברות למאורע הכולל שתי הצלחות בדיוק (ולכן כישלון אחד בדיוק) נקבל את ההסתברות $p^{2}\left(1-p\right)$. אם נחליט שהסדר לא משנה, נכפיל הסתברות זו במספר המאורעות האפשריים, שבמקרה של $n=3$, $k=2$ הוא $\left(\begin{array}{c} 3\\ 2 \end{array}\right)=3$, ולכן ההסתברות לשתי הצלחות בדיוק היא $3p^{2}\left(1-p\right)$. נניח כי X הוא משתנה מקרי שסופר את מספר ההצלחות. הערכים האפשריים של X הם כל השלמים בין 0 ל־n. נניח כי $n=5$, אז נקבל למשל את ההסתברות:

$$\mathrm{P}\left(X=2\right)=\left(\begin{array}{c} 5\\ 2 \end{array}\right)p^{2}\left(1-p\right)^{3}$$

כי $\binom{5}{2}$ הוא מספר האפשרויות לסידור של 2 הצלחות ב–5 ניסויים. p^2 היא ההסתברות ל–2 הצלחות. $(1-p)^3$ היא ההסתברות ל–3 כשלונות.
וכך נקבל את שאר ההסתברויות לשאר המקרים:

$$\mathrm{P}(X=0)=\binom{5}{0}p^0(1-p)^5=(1-p)^5$$

$$\mathrm{P}(X=1)=\binom{5}{1}p^1(1-p)^4=5p(1-p)^4$$

$$\mathrm{P}(X=2)=\binom{5}{2}p^2(1-p)^3=10p^2(1-p)^3$$

$$\mathrm{P}(X=3)=\binom{5}{3}p^3(1-p)^2=10p^3(1-p)^2$$

$$\mathrm{P}(X=4)=\binom{5}{4}p^4(1-p)^1=5p^4(1-p)$$

$$\mathrm{P}(X=5)=\binom{5}{5}p^5(1-p)^0=p^5$$

17.4 התפלגות גאומטרית

(משפחה חד־פרמטרית של התפלגויות)
מ`מ X ייקרא מפולג גאומטרית עם פרמטר p, $0<p\leq 1$, אם לכל $k\geq 1$ שלם פונקציית ההסתברות שלו היא:

$$\mathrm{P}(X=k)=p(1-p)^{k-1}\quad ,k\geq 1$$

מסמנים משתנה מקרי גאומטרי $X\sim Geo(p)$.

נוכיח שזו פונקציית התפלגות. ברור שהערכים המתקבלים הם חיוביים ממש. נוסחת הסכום של טור הנדסי אינסופי שאיברה הכללי הוא $a_n=aq^{n-1}$ כאשר $|q|<1$, היא:

$$\sum_{n=1}^{\infty}a_n=\frac{a}{1-q}$$

לכן במקרה הנוכחי מתקיים:

$$\sum_{k=1}^{\infty}\mathrm{P}(X=k)=\sum_{k=1}^{\infty}p(1-p)^{k-1}=\frac{p}{1-(1-p)}=1$$

הסבר להתפלגות הגאומטרית:

חוזרים באופן בלתי־תלוי על ניסוי בעל שתי תוצאות אפשריות, הצלחה וכישלון. ההסתברות להצלחה בכל ניסוי בודד היא p, כך שההסתברות המשלימה לכישלון היא $1-p$. המ`מ X סופר את מספר הניסויים עד וכולל ההצלחה הראשונה.

נבדוק מהי ההסתברות שחזרנו על הניסוי k פעמים בדיוק. מאורע זה משמעותו היא שנכשלנו $k-1$ פעמים ובפעם ה־ k הצלחנו, ולכן הסתברותו היא

$$\mathrm{P}\left(X=k\right)=\left(1-p\right)^{k-1}p$$

טענה: $P\left(X\geq k\right)=\left(1-p\right)^{k-1}$, $k\geq 1$. בפרט, $P\left(X=\infty\right)=0$

הוכחה:

$$\mathrm{P}\left(X\geq k\right)=\sum_{i=k}^{\infty}\mathrm{P}(X=i)=\sum_{i=k}^{\infty}p\left(1-p\right)^{i-1}=\frac{p\left(1-p\right)^{k-1}}{1-\left(1-p\right)}=\left(1-p\right)^{k-1}\overset{k\to\infty}{\longrightarrow}0$$

המסקנה מכך היא שבהסתברות 1, במוקדם או במאוחר, תופיע הצלחה.

17.5 התפלגות פואסון

(משפחה חד־פרמטרית של התפלגויות)

נזכיר: $e\cong 2.718...$ וכן: לכל x ממשי

$$\lim_{n\to\infty}\left(1+\tfrac{x}{n}\right)^{n}=e^{x}$$

$$\sum_{n=0}^{\infty}\tfrac{x^{n}}{n!}=1+x+\tfrac{x^{2}}{2}+\tfrac{x^{3}}{6}+...+\tfrac{x^{n}}{n!}+...=e^{x}$$

מ`מ X ייקרא מפולג פואסון עם פרמטר $\lambda>0$ אם לכל $k\geq 0$ שלם פונקציית ההסתברות שלו היא:

$$\mathrm{P}\left(X=k\right)=e^{-\lambda}\frac{\lambda^{k}}{k!}$$

מסמנים משתנה מקרי פואסוני $X\sim Pois\left(\lambda\right)$.

נוכיח שזו פונקציית התפלגות. ברור שהערכים המתקבלים הם חיוביים ממש.
נחשב את הסכום:

$$\sum_{k=0}^{\infty}\mathrm{P}\left(X=k\right)=\sum_{k=0}^{\infty}e^{-\lambda}\frac{\lambda^{k}}{k!}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^{k}}{k!}=e^{-\lambda}e^{\lambda}=1$$

הסבר להתפלגות פואסון:

נראה שהתפלגות פואסון היא קירוב להתפלגות הבינומית, כאשר $p \to 0$, $n \to \infty$ ו $np \to \lambda$.
נניח שנתון מ`מ $X \sim Bin\left(n,p\right)$. ניתן לראות שמתקיים:

$$\frac{\mathrm{P}\left(X=k+1\right)}{\mathrm{P}\left(X=k\right)}=\frac{\left(\begin{array}{c} n\\ k+1 \end{array}\right)}{\left(\begin{array}{c} n\\ k \end{array}\right)}\cdot\frac{p^{k+1}\left(1-p\right)^{n-k-1}}{p^{k}\left(1-p\right)^{n-k}}=\frac{n-k}{k+1}\cdot\frac{p}{1-p}$$

נסמן $\lambda=np$, ונשים לב שמתקיים:

$$\frac{\mathrm{P}\left(X=k+1\right)}{\mathrm{P}\left(X=k\right)}=\frac{n-k}{k+1}\cdot\frac{p}{1-p}=\frac{np-kp}{\left(k+1\right)\left(1-p\right)}\ \underset{\substack{n\to\infty\\ p\to0}}{\longrightarrow}\ \frac{\lambda}{k+1}$$

$$\Downarrow$$

ולכן, בגבול:

$$\mathrm{P}\left(X=k+1\right)\ \underset{\substack{n\to\infty\\ p\to0}}{\longrightarrow}\ \frac{\lambda}{k+1}\cdot\mathrm{P}\left(X=k\right)$$

מכאן

$$\mathrm{P}(X=k+1)=\frac{\lambda}{k+1}\cdot\frac{\lambda}{k}\cdot\mathrm{P}(X=k-1)$$

וכן הלאה, כך שבאינדוקציה נסיק כי מתקיים:

$$\mathrm{P}\left(X=k\right)=\frac{\lambda^{k}}{k!}\cdot\mathrm{P}(X=0),\ \ k\geq0$$

כל שנותר הוא למצוא את ערכו של $\mathrm{P}(X=0)$. ואכן,

$$1=\sum_{k=0}^{\infty}\mathrm{P}(X=k)=\sum_{k=0}^{\infty}\frac{\lambda^{k}}{k!}\cdot\mathrm{P}(X=0)=\mathrm{P}(X=0)\cdot\sum_{k=0}^{\infty}\frac{\lambda^{k}}{k!}=\mathrm{P}(X=0)\cdot e^{\lambda}$$

ולכן,

$$\mathrm{P}(X=0)=e^{-\lambda}$$

ואז

$$\mathrm{P}(X=k)=\frac{\lambda^k}{k!}\cdot e^{-\lambda}$$

ניתן לראות זאת גם בדרך הבאה:

$$\mathrm{P}\left(X=0\right)=\left(\begin{array}{c}n\\0\end{array}\right)p^0\left(1-p\right)^{n-0}=\left(1-p\right)^n=\left(1-\frac{\lambda}{n}\right)^n\underset{n\to\infty}{\longrightarrow}e^{-\lambda}$$

$$\mathrm{P}\left(X=1\right)=np\left(1-p\right)^{n-1}=\lambda\left(1-\tfrac{\lambda}{n}\right)^{n-1}\to\lambda e^{-\lambda}$$

$$\mathrm{P}\left(X=2\right)=\left(\begin{array}{c}n\\2\end{array}\right)p^2\left(1-p\right)^{n-2}=\tfrac{n(n-1)}{2}\cdot\left(\tfrac{\lambda}{n}\right)^2\left(1-\tfrac{\lambda}{n}\right)^{n-2}\to\tfrac{\lambda^2}{2}\cdot e^{-\lambda}$$

$$\vdots$$

כלומר, עבור n מספיק גדול, pמספיק קטן ו־np`בינוני`, ניתן לקרב באמצעות התפלגות פואסון את ההתפלגות הבינומית.

דוגמה: נניח כי $X\sim Bin\left(1000,0.002\right)$. נרצה לבדוק מהו ערכו של $\mathrm{P}\left(X=3\right)$. הערך המדויק בהתאם להגדרת ההתפלגות הבינומית הוא:

$$\mathrm{P}\left(X=3\right)=\left(\begin{array}{c}1000\\3\end{array}\right)0.002^3\cdot\left(1-0.002\right)^{997}=0.18062...$$

נשים לב שבמקרה זה $np=1000\cdot0.002=2$, ולכן לפי הקירוב הפואסוני נקבל:

$$\mathrm{P}\left(X=3\right)\cong e^{-2}\cdot\frac{2^3}{3!}=0.18044...$$

קיבלנו קירוב מדויק עד שלוש ספרות לאחר הנקודה.

כלומר, ככל ש-n גדול ו-p קטן, אין חשיבות לערכם המדויק של n,p ומספיק לדעת את מכפלתם כדי לקבל ערך מקורב להסתברות הבינומית המבוקשת מזו הפואסונית.

17.6 התפלגות בינומית שלילית

(משפחה דו־פרמטרית של התפלגויות)
מ`מ X ייקרא בעלת **התפלגות בינומית שלילית** עם הפרמטרים $0>p<1$, $r\geq1$ שלם, אם פונקציית ההסתברות שלו היא:

$$\mathrm{P}\left(X=k\right)=\left(\begin{array}{c}k-1\\r-1\end{array}\right)p^r\left(1-p\right)^{k-r}\ ,\ \ k\geq r$$

מסמנים משתנה מקרי מפולג בינומית שלילית $X\sim NB\left(r,p\right)$.
(לא נוכיח במסגרת זו שזו אכן פונקציית התפלגות.)

הסבר להתפלגות הבינומית השלילית:

נניח שמבצעים ניסוי עם שתי תוצאות אפשריות, הצלחה בהסתברות p וכישלון בהסתברות $1-p$. חוזרים על הניסוי באופן בלתי־תלוי עד וכולל ההצלחה ה־r.
נבדוק מהי ההסתברות שיידרשו בדיוק k ניסיונות כדי להגיע להצלחה ה־r?
ברור שהסתברות זו היא 0 עבור כל k המקיים $k<r$, כי לא ניתן להגיע ל־r הצלחות אם ביצענו פחות מ־r ניסויים. לכן נניח מעתה כי $k \geq r$.
נסמן ב־X את המ`מ שסופר את מספר הניסיונות.
הביטוי $p^r\left(1-p\right)^{k-r}$ הוא ההסתברות לכל סדרה באורך k שכוללת בדיוק k הצלחות.
כמה סדרות כאלו המסתיימות בהצלחה קיימות?
הניסיון האחרון, דהיינו הניסיון ה־k, הוא בהכרח הצלחה. לכן לצורך קיום התנאי נותר לבחור מתוך שאר $k-1$ הניסיונות את מיקומן של $r-1$ הצלחות, ולכן קיבלנו את הגודל $\left(\begin{array}{c} k-1 \\ r-1 \end{array}\right)$.

17.7 התפלגות היפר־גאומטרית

(משפחה תלת־פרמטרית של התפלגויות)
מ`מ X ייקרא בעלת **התפלגות היפר־גאומטרית** עם הפרמטרים n,a,b, כולם שלמים, כאשר $1 \leq n \leq a+b$, אם פונקציית ההסתברות שלו היא:

$$\mathrm{P}\left(X=k\right)=\frac{\left(\begin{array}{c} a \\ k \end{array}\right)\left(\begin{array}{c} b \\ n-k \end{array}\right)}{\left(\begin{array}{c} a+b \\ n \end{array}\right)} \quad ,\max\{0,n-b\} \leq k \leq \min\{a,n\}$$

מסמנים משתנה מקרי המפולג היפר־גאומטרית $X \sim HG\left(n,a,b\right)$.

הסבר להתפלגות ההיפר־גאומטרית:

נניח שנתונים a עצמים מסוג A ו־b עצמים מסוג B.
נגדיר את הבחירה בעצם מסוג A כ`הצלחה`, כך שההסתברות בבחירה אחת היא:

$$p=\frac{a}{a+b} \quad 1-p=\frac{b}{a+b}$$

נגדיר את המ`מ X להיות מספר העצמים מסוג A שנדגום.
נניח שמוציאים n עצמים באופן אקראי מתוך כלל העצמים עם החזרה, כמובן $X \sim Bin\left(n,\frac{a}{a+b}\right)$.
כעת נניח שמוציאים את העצמים ללא החזרה. כמובן מניחים $n \leq a+b$. יתר על כן $0 \leq x \leq \min\{a,n\}$ ו $0 \leq n-x \leq b$, ולכן, x יכול לקבל רק ערכים בין $\max\{0,n-b\}$ ובין $\min\{a,n\}$.

נראה שההסתברות להוציא בדיוק k עצמים מסוג A, עבור k בטווח זה, היא ההסתברות ההיפר־גאומטרית שהגדרנו.
המכנה מציין את מרחב המדגם, כלומר את מספר האפשרויות לבחור n עצמים מתוך $a+b$.

במונה, האיבר הראשון הוא מספר הדרכים לבחור k עצמים מתוך a, והאיבר השני הוא מספר הדרכים לבחור את שאר $n-k$ האיברים מתוך b האחרים. כופלים את שני האיברים כדי לקבל את שני האירועים יחד.
נשים לב כי אם התנאי הכפול $n-b \leq k \leq a$ לא מתקיים, ההסתברות הזו שווה ל-0.
נוכיח שזו אכן פונקציית הסתברות.
ראשית קל לראות שכל הערכים שמתקבלים חיוביים.
נחשב את סכום ההסתברויות, ולשם כך נשים לב שלאחר העברת אגפים מספיק להוכיח:

$$\sum_{k=0}^{n}\left(\begin{array}{c} a \\ k \end{array}\right)\left(\begin{array}{c} b \\ n-k \end{array}\right)=\left(\begin{array}{c} a+b \\ n \end{array}\right)$$

נזכור:

$$(x+1)^{a}(x+1)^{b}=(x+1)^{a+b}$$

ולכן נסיק:

$$(x+1)^{a}(x+1)^{b}=\sum_{i=0}^{a}\left(\begin{array}{c} a \\ i \end{array}\right)x^{i}\sum_{j=0}^{b}\left(\begin{array}{c} b \\ j \end{array}\right)x^{j}=\sum_{n=0}^{a+b}\left(\begin{array}{c} a+b \\ n \end{array}\right)x^{n}=(x+1)^{a+b}$$

מאחר והמקדמים של x^k בשני האגפים זהים, לכל $0 \leq k \leq a+b$, נסיק:

$$\sum_{i=0}^{n}\left(\begin{array}{c} a \\ i \end{array}\right)\left(\begin{array}{c} b \\ n-i \end{array}\right)=\left(\begin{array}{c} a+b \\ n \end{array}\right)$$

במחוברים שבהם $i>a$ או $n-i \geq b$ יש לקרוא את האיבר שבסכום השמאלי כ-0.

18 מדדי מרכז של משתנים מקריים

18.1 תוחלת של משתנה מקרי (Expected value)

הגדרה: יהי X מ`מ כלשהו עם פונקציית ההתפלגות $\mathrm{P}(X=x)$.
התוחלת של X מוגדרת ומסומנת להיות:

$$\mathrm{E}(X)=\sum_{x} x \cdot \mathrm{P}(X=x)$$

דוגמה: נניח כי $A \sim U[1,6]$, כך שפונקציית ההתפלגות היא:

X	1	2	3	4	5	6
$\mathrm{P}(X=x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

נחשב את התוחלת:

$$\mathrm{E}(X)=1 \cdot \frac{1}{6}+2 \cdot \frac{1}{6}+3 \cdot \frac{1}{6}+4 \cdot \frac{1}{6}+5 \cdot \frac{1}{6}+6 \cdot \frac{1}{6}=3.5$$

קיבלנו ערך שנמצא בדיוק באמצע שני ערכי הקיצון של X.
אם היינו משנים את המשקל של כל ערך, ומגדירים את פונקציית ההתפלגות כך:

X	1	2	3	4	5	6
$\mathrm{P}(X=x)$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{3}{12}$	$\frac{1}{6}$

היינו מקבלים את התוחלת הבאה:

$$\mathrm{E}(X)=1\cdot\frac{1}{6}+2\cdot\frac{1}{12}+3\cdot\frac{1}{6}+4\cdot\frac{1}{6}+5\cdot\frac{3}{12}+6\cdot\frac{1}{6}>3.5$$

כלומר, התוחלת גדלה כי העברנו משקל רב יותר לערך יחסית גדול.

התוחלת היא ערך שלוקח בחשבון לא רק את הערכים האפשריים אלא גם את המשקל של כל אחד מהם. כלומר את ההסתברות שכל אחד מהם יתרחש. התוחלת מהווה מעין סיכום כללי של המשתנה.
התוחלת מקיימת את התכונות שראינו לגבי הממוצע. (ראה עמוד 7) כך למשל היא משמרת את יחידות המדידה.

18.1.1 תוחלת של מ`מ ברנולי

$$X\sim B(p)\ ,\ \ \mathrm{E}(X)=p$$

x	0	1
$\mathrm{P}(X=x)$	$1-p$	p

הוכחה: נחשב לפי הגדרת התוחלת:

$$\mathrm{E}(X)=0\cdot(1-p)+1\cdot p=p$$

■

הערה: נשים לב שמתקיים $X\sim Bin(1,p) \Leftrightarrow X\sim B(p)$.

18.1.2 תוחלת של מ`מ אחיד

$$X\sim U(a,b)\ ,\ \ \mathrm{E}(X)=\frac{b+a}{2}$$

x	a	$a+1$	$a+2$	$\ldots$	$b-2$	$b-1$	b
$\mathrm{P}(X=x)$	$\frac{1}{b-a+1}$	$\frac{1}{b-a+1}$	$\frac{1}{b-a+1}$	$\ldots$	$\frac{1}{b-a+1}$	$\frac{1}{b-a+1}$	$\frac{1}{b-a+1}$

הוכחה: נחשב לפי הגדרת התוחלת:

$$\mathrm{E}\left(X\right)=\sum_{k=a}^{b}k\cdot\mathrm{P}\left(X=k\right)=\sum_{k=a}^{b}k\cdot\frac{1}{b-a+1}$$

$$=\frac{1}{b-a+1}\cdot\sum_{k=a}^{b}k=\frac{1}{b-a+1}\cdot\frac{(b-a+1)(a+b)}{2}=\frac{b+a}{2}$$

■

18.1.3 תוחלת של מ`מ בינומי

$$X\sim Bin\left(n,p\right)\ ,\ \ \mathrm{E}\left(X\right)=np$$

הוכחה: נחשב לפי הגדרת התוחלת (שימו לב לאינדקסים):

$$\mathrm{E}\left(X\right)=\sum_{k=0}^{n}k\cdot\mathrm{P}\left(X=k\right)=\sum_{k=0}^{n}k\cdot\frac{n!}{k!\left(n-k\right)!}\cdot p^{k}\left(1-p\right)^{n-k}$$

$$=np\cdot\sum_{k=1}^{n}\frac{\left(n-1\right)!}{\left(k-1\right)!\left(n-1-\left(k-1\right)\right)!}\cdot p^{k-1}\left(1-p\right)^{n-1-\left(k-1\right)}$$

$$=np\cdot\sum_{k=0}^{n-1}\frac{\left(n-1\right)!}{k!\left(n-1-k\right)!}\cdot p^{k}\left(1-p\right)^{n-1-k}=np$$

השוויון האחרון נובע מכך שהביטוי $\sum_{k=0}^{n-1}\frac{(n-1)!}{k!(n-1-k)!}\cdot p^{k}(1-p)^{n-1-k}$ הוא סכום ההסתברויות של מ`מ $Y\sim Bin\left(n-1,p\right)$, ולכן מהגדרת פונקציית התפלגות של מ`מ סכום זה שווה 1. ■

18.1.4 תוחלת של מ`מ פואסון

$$X\sim Pois\left(\lambda\right)\ ,\ \ \mathrm{E}\left(X\right)=\lambda$$

הוכחה: נחשב לפי הגדרת התוחלת:

$$\mathrm{E}\left(X\right)=\sum_{k=0}^{\infty}k\cdot\mathrm{P}\left(X=k\right)=\sum_{k=1}^{\infty}k\cdot e^{-\lambda}\cdot\frac{\lambda^{k}}{k!}$$

$$=\sum_{k=1}^{\infty}e^{-\lambda}\cdot\frac{\lambda^{k}}{\left(k-1\right)!}=\lambda\cdot\sum_{k=1}^{\infty}e^{-\lambda}\cdot\frac{\lambda^{k-1}}{\left(k-1\right)!}=\sum_{k=0}^{\infty}e^{-\lambda}\frac{\lambda^{k}}{k!}=\lambda$$

השוויון האחרון נובע מכך שהביטוי $\sum_{k=0}^{n} e^{-\lambda} \cdot \frac{\lambda^k}{k!}$ הוא סכום ההסתברויות של מ`מ $Y \sim Pois\left(\lambda\right)$, ולכן מהגדרת פונקציית התפלגות של סכום מ`מ סכום זה שווה ל־1.
■

18.1.5 תוחלת של מ`מ גאומטרי

$$X \sim Geo\left(p\right)\ ,\ \ \mathrm{E}\left(X\right) = \frac{1}{p}$$

הוכחה ראשונה: נחשב לפי הגדרת התוחלת:

$$\mathrm{E}\left(X\right) = \sum_{k=1}^{n} k \cdot \mathrm{P}\left(X = k\right) = \sum_{k=1}^{n} k \cdot p\left(1-p\right)^{k-1} = 0 + p + 2p\left(1-p\right) + 3p\left(1-p\right)^2 + ...$$
$$\Downarrow$$
$$\left(1-p\right)\mathrm{E}\left(X\right) = \left(1-p\right)\left(p + 2p\left(1-p\right) + 3p\left(1-p\right)^2 + ...\right)$$
$$\Downarrow$$
$$\left(1-p\right)\mathrm{E}\left(X\right) = p\left(1-p\right) + 2p\left(1-p\right)^2 + 3p\left(1-p\right)^3 ...$$
$$\Downarrow$$
$$p \cdot \mathrm{E}\left(X\right) = \mathrm{E}\left(X\right) - \left(p\left(1-p\right) + 2p\left(1-p\right)^2 + 3p\left(1-p\right)^3 ...\right)$$
$$\Downarrow$$
$$p \cdot \mathrm{E}\left(X\right) = \left(p + 2p\left(1-p\right) + 3p\left(1-p\right)^2 + ...\right) - \left(p\left(1-p\right) + 2p\left(1-p\right)^2 + 3p\left(1-p\right)^3 ...\right)$$
$$\Downarrow$$
$$p \cdot \mathrm{E}\left(X\right) = p + p\left(1-p\right) + p\left(1-p\right)^2 + p\left(1-p\right)^3 + ... = 1$$
$$\Downarrow$$
$$\mathrm{E}\left(X\right) = \tfrac{1}{p}$$

השוויון הלפני־אחרון נובע מכך שהביטוי $\sum_{k=1}^{\infty} p(1-p)^{k-1}$ הוא סכום ההסתברויות של מ`מ $Y \sim Geo\left(\lambda\right)$, ולכן מהגדרת התפלגות של סכום מ`מ סכום זה שווה ל־1. ■

- נוכיח באופן נוסף שזו התוחלת של מ`מ גאומטרי, ולצורך כך נוכיח טענת־עזר.

טענת־עזר: אם X מ`מ שלם ואי־שלילי, אז:

$$\mathrm{E}\left(X\right) = \sum_{k=1}^{\infty} \mathrm{P}\left(X \geq k\right)$$

הוכחה: נסמן $P\left(X = x\right) = p_x$ ונקבל לפי הגדרת התוחלת:

$$\mathrm{E}\left(X\right) = \sum_{k=1}^{\infty} k \cdot \mathrm{P}\left(X = k\right) = p_1 + 2p_2 + 3p_3 + 4p_4 + ...$$

נתבונן בסכום שקיבלנו באופן הבא:

$$\begin{array}{lccccccccc} \mathrm{P}(X\geq 1) = & p_1 & + & p_2 & + & p_3 & + & p_4 & +\ldots \\ \mathrm{P}(X\geq 2) = & & & p_2 & + & +p_3 & + & p_4 & +\ldots \\ \mathrm{P}(X\geq 3) = & & & & & p_3 & + & p_4 & +\ldots \\ \mathrm{P}(X\geq 4) = & & & & & & & p_4 & +\ldots \\ \vdots & & & & & & & & \ddots \end{array}$$

נשים לב שזה בדיוק אותו סכום שהצגנו לעיל לפי הגדרת התוחלת, ולכן התוחלת ניתנת להצגה כסכום $\sum_{k=1}^{\infty}\mathrm{P}(X\geq k)$.

הוכחה שנייה: ראינו לעיל בהתפלגות הגאומטרית, שמתקיים $\mathrm{P}(X\geq k)=(1-p)^{k-1}$, עבור $k\geq 1$.
נסיק:

$$\mathrm{E}(X)=\sum_{k=1}^{\infty}\mathrm{P}(X\geq k)=\sum_{k=1}^{\infty}(1-p)^{k-1}=\frac{1}{1-(1-p)}=\frac{1}{p}$$

השוויון השלישי נובע מנוסחת הסוכם של טור הנדסי, שאיברו הראשון הוא 1 ומנת הטור היא $1-p$. ■

18.1.6 תוחלת של מ`מ בינומי שלילי

$$X\sim NB(r,p)\quad \mathrm{E}(X)=\frac{r}{p}$$

הוכחה לכך נראה בהמשך - עמוד 121

18.1.7 תוחלת של מ`מ היפר־גאומטרי

$$X\sim HG(n,a,b)\quad \mathrm{E}(X)=n\cdot\frac{a}{a+b}$$

הוכחה לכך נראה בהמשך - עמוד 120

18.2 שכיח

הגדרה: השכיח של מ`מ הוא הערך x שמקבל את ההסתברות הגבוהה ביותר. כלומר $\arg\max_{x}\{\mathrm{P}(X=x)\}$.

18.2.1 שכיח של מ`מ ברנולי

נניח כי $X\sim Bin\left(1,p\right)$ (למעשה זהו מ`מ ברנולי), נקבל:

$$\text{Mode}=\begin{cases} 1 & p>\frac{1}{2}\\ 0 & p<\frac{1}{2} \end{cases}$$

במקרה $p=\frac{1}{2}$ שני הערכים שכיחים באותה מידה ולכן שניהם שכיחים.

18.2.2 שכיח של מ`מ פואסון

נוכיח שהשכיח של מ`מ $X\sim Pois\left(\lambda\right)$ הוא $\lfloor\lambda\rfloor$ (הערך השלם).
נתון כי $k\geq0$ וכן נשים לב לקשר שבין p_{k} ל-p_{k-1}:

$$p_{k}=e^{-\lambda}\cdot\frac{\lambda^{k}}{k!}=e^{-\lambda}\cdot\frac{\lambda^{k-1}}{\left(k-1\right)!}\cdot\frac{\lambda}{k}=p_{k-1}\cdot\frac{\lambda}{k}$$

נסיק שכאשר $\frac{\lambda}{k}>1$ אז $p_{k}>p_{k-1}$, וכאשר $\frac{\lambda}{k}<1$ אז $p_{k}<p_{k-1}$.
מכאן ש-p_{k} כפונקציה של k, בתחילה עולה ואח`כ יורדת.
השכיח הוא ה-k האחרון שעבורו $p_{k-1}<p_{k}$. כלומר ה-k האחרון שעבורו $k<\lambda$, משמע הערך השלם של λ. ■

18.2.3 שכיח של מ`מ בינומי

נוכיח שהשכיח של מ`מ $X\sim Bin\left(n,p\right)$ הוא $\lfloor\left(n+1\right)p\rfloor$.
נשים לב לקשר שבין p_{k} ל-p_{k-1}:

$$p_{k}=\tfrac{n!}{k!\left(n-k\right)!}\cdot p^{k}\left(1-p\right)^{n-k}$$

$$=\tfrac{n!}{\left(k-1\right)!\left(n-k+1\right)!}\cdot p^{k-1}\left(1-p\right)^{n-k+1}\cdot\tfrac{\left(n-k+1\right)p}{k\left(1-p\right)}=p_{k-1}\cdot\tfrac{\left(n-k+1\right)p}{k\left(1-p\right)}$$

כמו־כן מתקיים:

$$\left(n+1\right)p>k\Leftrightarrow\left(n-k+1\right)p>k\left(1-p\right)$$

ולכן נקבל שכאשר $\left(n+1\right)p>k$ אז $p_{k}>p_{k-1}$, וכאשר $\left(n+1\right)p<k$ אז $p_{k}<p_{k-1}$.
מכאן ש-p_{k} כפונקציה של k, בתחילה עולה ואח`כ יורדת (פונקציה יונימודלית).
השכיח הוא ה-k האחרון שעבורו $p_{k-1}<p_{k}$. כלומר ה-k האחרון שעבורו $\left(n+1\right)p>k$, משמע הערך השלם של $\left(n+1\right)p$. ■

18.2.4 שכיח של מ`מ גאומטרי

מ`מ גאומטרי הוא דוגמה למקרה בו משמעותו של השכיח כמדד מרכזי היא בעייתית.
נוכיח שעבור מ`מ $X\sim Geo\left(p\right)$ השכיח הוא 1, ללא כל תלות ב-p.
נשים לב שמתקיים $p_{k}=p\left(1-p\right)^{k-1}$.
פונקציה זו מונוטונית יורדת ב-k, ולכן הערך המקסימלי מתקבל עבור 1, ללא קשר לערכו של p.

18.3 תוחלת של פונקציות של משתנים מקריים

נניח כי X מ`מ כלשהו, ונניח כי f היא פונקציה של X.
כלומר דוגמים מ`מ, מפעילים עליו את f ומתייחסים לתוצאה שהתקבלה, שנסמנה $f\left(X\right)$.
נשים לב שהערך שמתקבל הוא עצמו מ`מ בעל התפלגות ותוחלת.

דוגמה: נניח כי $X\sim U\left[1,6\right]$, $f\left(x\right)=x^{2}$.
בחישוב פשוט נקבל את ההתפלגות:

x	1	2	3	4	5	6
X^{2}	1	4	9	16	25	36
$\mathrm{P}\left(X^{2}=x^{2}\right)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

וכן חישוב נוסף ייתן לנו את התוחלת של המ`מ X^{2}:

$$\mathrm{E}\left(X^{2}\right)=\frac{1}{6}\left(1+4+9+16+25+36\right)=\frac{91}{6}=15\frac{1}{6}$$

הגדרה: תוחלת של פונקציה של מ`מ היא:

$$\boxed{\mathrm{E}\left[f\left(X\right)\right]=\sum_{x}f\left(x\right)\cdot\mathrm{P}\left(X=x\right)}$$

18.3.1 תוחלת של פונקציה־לינארית

נניח כי f היא פונקציה לינארית. כלומר היא מהצורה $f\left(x\right)=a+bx$.

טענה: התוחלת של פונקציה מהצורה הזו היא:

$$\boxed{\mathrm{E}\left[a+bX\right]=a+b\mathrm{E}\left(X\right)}$$

כלומר, עבור פונקציות לינאריות, תוחלת הפונקציה שווה לפונקציה של התוחלת (כפי שגם מתקיים בממוצעים).
תובנה זו מאפשרת לעבור בין יחידות מדידה שונות, שכן מדובר בפונקציה לינארית.

הוכחה: נחשב לפי הגדרת התוחלת:

$$\mathrm{E}\left[f\left(X\right)\right]=\mathrm{E}\left(a+bX\right)=\sum_{x}\left(a+bx\right)\mathrm{P}\left(X=x\right)$$

$$=a\sum_{x}\mathrm{P}\left(X=x\right)+b\sum_{x}x\mathrm{P}\left(X=x\right)=a+b\cdot\mathrm{E}\left(X\right)$$

■

דוגמה: נמשיך את הדיון בדוגמה $X\sim U\left[1,6\right]$, $f\left(x\right)=x^2$, ונקבל שמכיוון שזו אינה פונקציה לינארית, אז:

$$\mathrm{E}\left[f\left(X\right)\right]=\mathrm{E}\left(X^2\right)=\frac{91}{6}>\mathrm{E}^2\left(X\right)=\left(\frac{7}{2}\right)^2$$

בהמשך נוכיח שלא מדובר במקרה, ולכל מ`מ X מתקיים $\mathrm{E}^2\left(X\right)\leq\mathrm{E}\left(X^2\right)$, באופן אנלוגי לתוצאה שהראינו בפרק על סטטיסטיקה תאורית, $\overline{X}^2\leq\overline{X^2}$.

18.3.2 תוחלת של הרכבת פונקציות

נניח שנתונות שתי פונקציות $f\left(x\right),g\left(x\right)$. נגדיר פונקציה שלישית שמהווה פעולה לינארית על הפונקציות הללו.
כלומר ניקח את הפונקציה הלינארית $h\left(y,z\right)=ay+bz+c$ ונרכיב אותה על הפונקציות הללו שמופעלות על המ`מ:

$$h\left(f\left(X\right),g\left(X\right)\right)=af\left(X\right)+bg\left(X\right)+c$$

טענה: התוחלת של ההרכבה היא:

$$\mathrm{E}\left[h\left(f\left(X\right),g\left(X\right)\right)\right]=a\mathrm{E}\left[f\left(X\right)\right]+b\mathrm{E}\left[g\left(X\right)\right]+c$$

הוכחה:

$$\mathrm{E}\left[h\left(f\left(X\right),g\left(X\right)\right)\right]=\sum_x h\left(f\left(X\right),g\left(X\right)\right)\mathrm{P}\left(X=x\right)$$

$$=\sum_x\left(af\left(X\right)+bg\left(X\right)+c\right)\cdot\mathrm{P}\left(X=x\right)$$

$$=a\sum_x f(X)\mathrm{P}\left(X=x\right)+b\sum_x g\left(X\right)\mathrm{P}\left(X=x\right)+c\cdot\sum_x\mathrm{P}\left(X=x\right)$$

$$=a\mathrm{E}\left[f\left(X\right)\right]+b\mathrm{E}\left[g\left(X\right)\right]+c$$

■

חשוב לשים לה שהטענה אינה נכונה עבור מכפלה. כלומר:

$$\mathrm{E}\left[f\left(X\right)\cdot g\left(X\right)\right]\neq\mathrm{E}\left[f\left(X\right)\right]E\left[g\left(X\right)\right]$$

19 מדדי פיזור של משתנים מקריים

19.1 שונות של משתנים מקריים

בדומה למדד הפיזור עבור התצפיות שדנו בו בפרק של סטטיסטיקה תיאורית, נציע מדד פיזור למשתנים מקריים.

הגדרה: השונות של מ`מ X מוגדרת ומסומנת:

$$\mathrm{Var}\left(X\right)=\mathrm{E}\left[X-\mathrm{E}\left(X\right)\right]^{2}$$

התוחלת $\mathrm{E}\left(X\right)$ קבועה בהינתן המ`מ.

נשים לב שהשונות היא בתוחלת של הפונקציה $f\left(X\right)=\left(X-\mathrm{E}\left(X\right)\right)^{2}$.

הסבר: אם כן, בשונות של משתנים מקריים אנו מסתכלים על המרחק של המ`מ ממדד מרכזי – התוחלת, כאשר את המרחק מגדירים כריבוע ההפרש.
נרצה מדד מרכזי למרחקים אלו, ולכן נבחר את התוחלת. כך קיבלנו את הגדרת השונות.

הסיבה שמסתכלים דווקא על ריבוע ההפרש, היא כי לו היינו מסתכלים למשל רק על ההפרש, אז תוחלת ההפרש הייתה אפס. כי אם נסמן $f\left(x\right)=x-\mathrm{E}\left(X\right)$, נקבל:

$$\mathrm{E}\left[X-\mathrm{E}\left(X\right)\right]=\mathrm{E}\left[f\left(X\right)\right]=f\left(\mathrm{E}\left(X\right)\right)=\mathrm{E}\left(X\right)-\mathrm{E}\left(X\right)=0$$

דוגמה: נניח כי $X\sim U\left[1,6\right]$. בחישוב מפורט שביצענו נבע כי התוחלת היא $\mathrm{E}\left(X\right)=\frac{7}{2}$.

x	1	2	3	4	5	6
$\mathrm{P}\left(X=x\right)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$\left(x-\frac{7}{2}\right)^{2}$	$\frac{25}{4}$	$\frac{9}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{9}{4}$	$\frac{25}{4}$

ומכאן שהשונות היא:

$$\mathrm{Var}\left(X\right)=\frac{1}{6}\cdot\frac{25}{4}+\frac{1}{6}\cdot\frac{9}{4}+\frac{1}{6}\cdot\frac{1}{4}+\frac{1}{6}\cdot\frac{9}{4}+\frac{1}{6}\cdot\frac{25}{4}=\frac{35}{12}$$

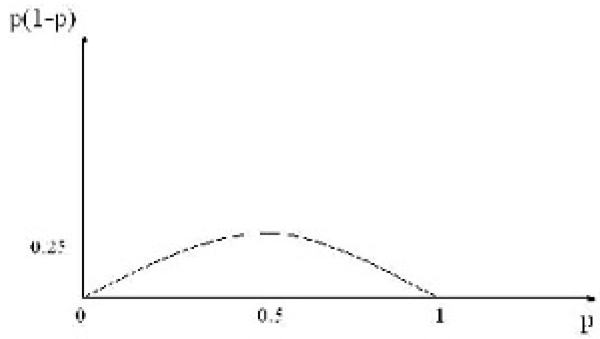
דוגמה: נניח כי $X\sim Bin\left(1,p\right)$. ראינו שהתוחלת של מ`מ ברנולי היא p. מכאן שהשונות היא $\mathrm{Var}\left(X\right)=\mathrm{E}\left[X-p\right]^{2}$:

x	0	1
$\mathrm{P}\left(X=x\right)$	$1-p$	p
$\left(x-p\right)^{2}$	p^{2}	$\left(1-p\right)^{2}$

ומכאן שהשונות היא:

$$\mathrm{Var}\left(X\right)=\left(1-p\right)p^{2}+p\left(1-p\right)^{2}=p\left(1-p\right)\left(p+1-p\right)=p\left(1-p\right)$$

הערה: השונות של מ`מ ברנולי היא $p(1-p)=p-p^2$. זו פונקציה ריבועית שסימטרית סביב $p=0.5$.
הגרף שלה נראה כך:



נשים לב כי השונות שווה לאפס כאשר p שווה 0 או 1. כלומר, התוצאה של המשתנה המקרי קבועה כ`כשלון` או `הצלחה` (בהתאמה).
השונות מקבלת ערך מקסימלי כאשר $p=0.5$, מה שמרמז על כך שכאשר $p=0.5$ מידת האי ודאות הקיימת במ`מ ברנולי היא מקסימלית.

מתכונות השונות

1. $\mathrm{Var}(X)\geq 0$
2. $\mathrm{Var}(X)=0$ אם ורק אם $\mathrm{P}(X=\mathrm{E}(X))=1$
כלומר קיים a קבוע, שעבורו $\mathrm{P}(X=a)=1$ (מ`מ דטרמיניסטי/מנוון).
3. יחידות המדידה של $\mathrm{Var}(X)$ הן ריבוע יחידות המדידה של X.
4. $\mathrm{Var}(X+a)=\mathrm{Var}(X)$
5. $\mathrm{Var}(aX)=a^2\mathrm{Var}(X)$

הערה:

$$\mathrm{Var}(f(X)+g(X))\neq\mathrm{Var}(f(X))+\mathrm{Var}(g(X))$$

19.1.1 סטיית תקן של משתנה מקרי

הגדרה: סטיית תקן של מ`מ X היא שורש השונות. מסמנים:

$$\mathrm{SD}\left(X\right)=\sqrt{\mathrm{Var}\left(X\right)}$$

תכונות סטיית התקן

1. $\mathrm{SD}\left(X\right)\geq 0$
2. $\mathrm{SD}\left(X\right)=0$ אם ורק אם $\mathrm{P}\left(X=\mathrm{E}\left(X\right)\right)=1$
כלומר קיים a קבוע, שעבורו $\mathrm{P}\left(X=a\right)=1$ (מ`מ דטרמיניסטי/מנוון).
3. יחידות המדידה של $\mathrm{SD}\left(X\right)$ הן אותן יחידות המדידה של X.
זוהי הסיבה שסטיית התקן היא מדד הפיזור המועדף.
4. $\mathrm{SD}\left(X+a\right)=\mathrm{SD}\left(X\right)$
5. $\mathrm{SD}\left(aX\right)=\left|a\right|\mathrm{SD}\left(X\right)$

19.1.2 נוסחה לחישוב השונות

$$\boxed{\mathrm{Var}\left(X\right)=\mathrm{E}\left(X^{2}\right)-\left[\mathrm{E}\left(X\right)\right]^{2}}$$

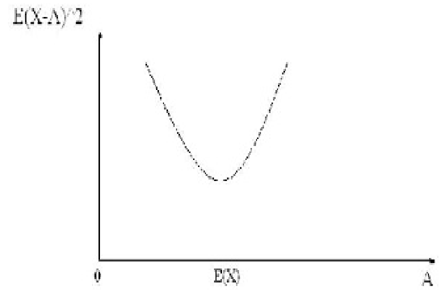
הוכחה: נשים לב שבאופן כללי מהנוסחה לפולינום ריבועי נובע:

$$\mathrm{E}\left(\left(X-A\right)^{2}\right)=\mathrm{E}\left(X^{2}-2XA+A^{2}\right)=\mathrm{E}\left(X^{2}\right)-2A\mathrm{E}\left(X\right)+A^{2}$$

נציב $A=\mathrm{E}(X)$ באגף ימין ונקבל

$$\mathrm{E}\left(X^{2}\right)-2\mathrm{E}\left(X\right)\mathrm{E}\left(X\right)+\mathrm{E}^{2}\left(X\right)=\mathrm{E}\left(X^{2}\right)-\mathrm{E}^{2}\left(X\right)$$

כנדרש.



המינימום של הפרבולה מתקבל עבור $\mathrm{E}(X) = -\frac{-2\mathrm{E}(X)}{2\cdot 1} = A_{\min}$.
כאשר $y = ax^2 + bx + c$ נקבל $y_{\min} = c - \frac{b^2}{4a}$.
לכן:

$$\mathrm{Var}(X) = \mathrm{E}(X - \mathrm{E}(X))^2 = \mathrm{E}(X-A)^2_{\min}$$

$$= \mathrm{E}\left(X^2\right) - \frac{[-2\mathrm{E}(X)]^2}{4} = \mathrm{E}\left(X^2\right) - [\mathrm{E}(X)]^2$$

■

מסקנה: ידוע כי $\mathrm{Var}(X) \geq 0$, ולכן נסיק:

$$0 \leq \mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}^2(X)$$
$$\Downarrow$$
$$\mathrm{E}^2(X) \leq \mathrm{E}\left(X^2\right)$$

19.1.3 שונות של מ`מ ברנולי

$$X \sim Bin(1,p)\ ,\ \mathrm{Var}(X) = p(1-p)$$

הוכחה: נניח כי $X \sim Bin(1,p)$ כך שההתפלגות שלו היא:

x	0	1
$\mathrm{P}(X=x)$	$1-p$	p
x^2	0	1

ולכן נסיק כי $\mathrm{E}\left(X^2\right)=\mathrm{E}\left(X\right)=p$, ונקבל:

$$\mathrm{Var}\left(X\right)=\mathrm{E}\left(X^2\right)-\mathrm{E}^2\left(X\right)=p-p^2=p\left(1-p\right)$$

■

19.1.4 שונות של מ`מ פואסון

$$X\sim Pois\left(\lambda\right)\ \ ,\ \ \mathrm{Var}\left(X\right)=\lambda$$

הוכחה: ראשית נוכיח את השוויון $\mathrm{E}\left(X^2\right)=\lambda^2+\lambda$ (שימו לב לאינדקסים):

$$\mathrm{E}\left(X^2\right)=\sum_{k=0}^{\infty}k^2e^{-\lambda}\cdot\frac{\lambda^k}{k!}=\sum_{k=0}^{\infty}k\left(k-1\right)e^{-\lambda}\cdot\frac{\lambda^k}{k!}+\sum_{k=0}^{\infty}ke^{-\lambda}\cdot\frac{\lambda^k}{k!}$$

$$=\lambda^2\sum_{k=2}^{\infty}\frac{e^{-\lambda}\lambda^{k-2}}{(k-2)!}+\lambda=\lambda^2\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}+\lambda=\lambda^2+\lambda$$

השוויון השלישי מבוסס על כך ש-$\mathrm{E}\left(X\right)=\lambda$.
כעת נוכל להסיק:

$$\mathrm{Var}\left(X\right)=\mathrm{E}\left(X^2\right)-\mathrm{E}^2\left(X\right)=\lambda^2+\lambda-\lambda^2=\lambda$$

■

19.1.5 שונות של מ`מ בינומי

$$X\sim Bin\left(n,p\right)\ \ ,\ \ \mathrm{Var}\left(X\right)=np\left(1-p\right)$$

נוכיח טענה זו בהמשך (עמוד 133)

19.1.6 שונות של מ`מ גאומטרי

$$X\sim Geo\left(p\right),\quad \mathrm{Var}\left(X\right)=\frac{1-p}{p^2}$$

הוכחה: הוכחנו כי ניתן להציג את התוחלת כסכום הבא:

$$\mathrm{E}\left(X^2\right)=\sum_{k=1}^{\infty}\mathrm{P}(X^2\geq k)=\sum_{k=1}^{\infty}\mathrm{P}(X\geq\lceil\sqrt{k}\rceil)$$
$$=\sum_{n=1}^{\infty}\sum_{k:\lceil\sqrt{k}\rceil=n}^{\infty}\mathrm{P}(X\geq n)$$
$$=\sum_{n=1}^{\infty}[n^2-(n-1)^2]\mathrm{P}(X\geq n)$$
$$=\sum_{n=1}^{\infty}(2n-1)(1-p)^{n-1}=\frac{2}{p^2}-\frac{1}{p}$$

השוויון האחרון מקורו בכך שמספר ערכי k כך ש $n=\lceil\sqrt{k}\rceil$ הוא $n^2-(n-1)^2$ (כדאי לבדוק זאת על ערכי kקטנים)■

ולכן

$$\mathrm{Var}\left(X\right)=\frac{2}{p^2}-\frac{1}{p}-(\frac{1}{p})^2=\frac{1-p}{p^2}$$

19.1.7 שונות של מ`מ אחיד

$$X\sim U\left[a,b\right],\quad \mathrm{Var}\left(X\right)=\frac{\left(b-a+2\right)\left(b-a\right)}{12}$$

הוכחה: ראשית נניח כי $a=1,b=n$. נקבל:

$$\mathrm{P}\left(X=i\right)=\frac{1}{n},\ 1\leq i\leq n$$
$$\mathrm{E}\left(X\right)=\frac{1+n}{2}$$
$$\mathrm{E}\left(X^2\right)=\frac{1}{n}\cdot1+\frac{1}{n}\cdot4+\frac{1}{n}\cdot16+...+\frac{1}{n}\cdot n^2=\frac{1}{n}\sum_{i=1}^{n}i^2$$
$$=\frac{1}{n}\cdot\frac{n(n+1)(2n+1)}{6}=\frac{(n+1)(2n+1)}{6}$$

השוויון השלישי ניתן להוכחה באמצעות אינדוקציה.
נסיק מכאן כי השונות היא:

$$\mathrm{Var}\left(X\right)=\mathrm{E}\left(X^2\right)-\mathrm{E}^2\left(X\right)=\frac{(n+1)(2n+1)}{6}-\frac{(n+1)^2}{4}$$
$$=\frac{(n+1)(4n+2-3n-3)}{12}=\frac{(n+1)(n-1)}{12}=\frac{n^2-1}{12}$$

נסיק מכאן באופן כללי שאם $X\sim U\left[a,b\right]$, אז:

$$Y=X-a+1\sim U\left[1,b-a+1\right]$$

וכפי שהוכחנו מתקיים:

$$\mathrm{Var}\left(X-a+1\right)=\mathrm{Var}\left(X\right)$$

ומכיוון שכעת נמצא $b-a+1$ בתפקיד n, נקבל:

$$\mathrm{Var}\left(Y\right)=\mathrm{Var}\left(X\right)=\tfrac{(b-a+1)^2-1}{12}=$$

$$=\tfrac{(b-a)^2+2(b-a)+1-1}{12}=\tfrac{(b-a)(b-a+2)}{12}$$

■

19.2 פרדוקס המהמר (או: פרדוקס סנט־פטרבורג)

מהמר משחק בקזינו תחת הכללים הפשוטים הבאים: כל פעם הוא מהמר על סכום מסוים. אם הוא זוכה הוא מוסיף למאזנו סכום הכפול להשקעתו, ואם לא, הוא מפסיד את הכסף שהימר עליו.
נסמן ב־p את ההסתברות לזכייה. קשה לחשוב על מצב שבו $p>0.5$, כי אז תוחלת הרווח חיובית והקזינו יפסיד. נניח לכן כי $p\leq 0.5$.

מהמר חושב שמצא דרך להכות את הקזינו: הוא יתחיל להמר על דולר אחד. אם הוא ירוויח הוא יפרוש והרווח שלו הוא 1, אם יפסיד יהמר שוב, והפעם על 2 דולר. אם הוא ירוויח הוא יפרוש והרווח שלו הוא $2-1=1$. אם יפסיד יהמר שוב, והפעם על 4 דולר. אם הוא ירוויח הוא יפרוש והרווח שלו הוא $4-2-1=1$, וכן הלאה. כך מובטחת לו זכייה בסכום זעום של דולרים אמנם, אבל זכותו מובטחת. אולי יחזור על כך מספר פעמים.

בקורס בשנה א' למד המהמר שמספר הסיבובים עד לזכייה הראשונה מפולג גאומטרית עם פרמטר p, וכן שסכום ההסתברויות הוא 1. לכן במוקדם או במאוחר הוא יזכה ויפרוש לביתו עם רווח של דולר בכיסו. האמנם?

הסבר לפרדוקס: לרשות המהמר חייב להיות סכום התחלתי סופי, ואם יפסיד את כולו הוא לא יוכל להמשיך לשחק. סכום זה יכול להיות גדול (אם למשל הוא מכר את ביתו לצורך כך), אבל הוא בכל מקרה סופי
מכאן שכדי שהמהר יוכל לשחק n משחקים, אם יהיה צורך בכך, הסכום ההתחלתי שחייב להיות ברשותו (לפי נוסחת הסכום של טור הנדסי) הוא:

$$1+2+...+2^n=1\cdot\frac{2^n-1}{2-1}=2^n-1$$

נניח שהמהר מחזיק באמתחתו את הסכום הזה, ובסוף הערב המהמר מרוויח דולר אחד, בהתאם לתכנית שלו, או מפסיד את כל כספו.
נניח לצורך הפשטות כי $p=0.5$ (ההסתברות האידאלית מבחינת המהמר), אז נקבל את פונקציית ההסתברות הבאה:

X	$\mathrm{P}\left(X=x\right)$
$-\left(2^n-1\right)$	$\left(\frac{1}{2}\right)^n$
1	$1-\left(\frac{1}{2}\right)^n$

נחשב את תוחלת הרווח של המהמר:

$$\mathrm{E}\left(X\right)=1\cdot\left(1-\left(\frac{1}{n}\right)^{n}\right)-\left(2^{n}-1\right)\left(\frac{1}{2}\right)^{n}=1-\left(\frac{1}{2}\right)^{n}-2^{n}\left(\frac{1}{2}\right)^{n}+\left(\frac{1}{2}\right)^{n}=0$$

כעת בדקו בעצמכם מדוע כאשר $p<0.5$ התוחלת הופכת להיות שלילית.

19.2.1 הערה: הסתברות ושכיחות יחסית

נניח כי $X\sim Bin\left(n,p\right)$ כך ש־$\mathrm{E}\left(X\right)=np$, $\mathrm{Var}\left(X\right)=np\left(1-p\right)$.
מבצעים ניסוי n פעמים באופן בלתי־תלוי, כאשר בכל פעם התוצאה A מוגדרת כ`הצלחה`, ונסמן $\mathrm{P}\left(A\right)=p$. נניח כי X הוא מספר ההצלחות. השכיחות היחסית של מספר ההצלחות היא כמובן $\frac{X}{n}$.
נשים לב שהשכיחות היחסית עצמה היא גם משתנה מקרי.
אם כן נבדוק מהי תוחלת השכיחות היחסית של מספר ההצלחות:

$$\mathrm{E}\left(\frac{X}{n}\right)=\frac{\mathrm{E}\left(X\right)}{n}=\frac{np}{n}=p$$

וכן נבדוק מהי השונות של השכיחות היחסית:

$$\mathrm{Var}\left(\frac{X}{n}\right)=\frac{1}{n^{2}}\cdot\mathrm{Var}\left(X\right)=\frac{1}{n^{2}}\cdot np\left(1-p\right)=\frac{p\left(1-p\right)}{n}\underset{n\to\infty}{\longrightarrow}0$$

כלומר, ככל שמספר הניסויים גדל, כך השונות של השכיחות היחסית קטנה.
התוחלת של השכיחות היחסית היא p (לכל n) ולכן השכיחות היחסית היא אמנם מקרית, אולם היא הולכת ומתקרבת להסתברות p ככל שמספר החזרות על הניסוי גדל, במובן שהשונות שלה שואפת ל־0.

19.2.2 הערה: סופיות התוחלת/השונות

נניח כי $X\sim Geo\left(\frac{1}{2}\right)$, ונניח $Y=\alpha^{X}$ עבור $\alpha>0$ כלשהו.
נתבונן בתוחלת של Y ושל Y^{2}:

$$\mathrm{E}\left(Y\right)=\sum_{i=1}^{\infty}\alpha^{i}\left(\tfrac{1}{2}\right)^{i}=\begin{cases}\frac{\frac{\alpha}{2}}{1-\frac{\alpha}{2}} & 0\le\alpha\le2\\ \infty & 2\le\alpha\end{cases}$$

$$\mathrm{E}\left(Y^{2}\right)=\sum_{i=1}^{\infty}\alpha^{2i}\left(\tfrac{1}{2}\right)^{i}=\begin{cases}\frac{\frac{\alpha^{2}}{2}}{1-\frac{\alpha^{2}}{2}} & 0\le\alpha\le\sqrt{2}\\ \infty & \sqrt{2}\le\alpha\end{cases}$$

מכאן שעבור המ`מ Y התוחלת סופית רק אם $0\le\alpha\le2$.
כדי שהשונות תהיה מוגדרת נצטרך לדרוש גם $\mathrm{E}\left(Y^{2}\right)<\infty$, ולכן השונות סופית רק אם $0\le\alpha\le\sqrt{2}$.

19.3 חציון

נניח שנתונה פונקציית הסתברות $\mathrm{P}\left(X=x\right)$, ונתונים הערכים:

$$x_1<x_2<x_3<...<x_n$$

נגדיר את ההסתברויות המתאימות הבאות:

$$p_1,p_2,p_3,...,p_n$$

כלומר $\mathrm{P}\left(X=x_i\right)=p_i$, $1\leq i\leq n$.

נגדיר פונקציית קנס $f\left(A\right)=\mathrm{E}\left(\left|X-A\right|\right)$, ונחפש ערך של A שימזער אותה.
נשים לב שלפי הגדרת התוחלת מתקיים:

$$f\left(A\right)=\sum_{i=1}^{n}\left|x_i-A\right|\mathrm{P}\left(X=x_i\right)=\sum_{i=1}^{n}\left|x_i-A\right|p_i$$

זוהי פונקציה לינארית למקוטעין ב־A שמזכירה פרבולה צוחקת. היא מקבלת מינימום כאשר הגרף משנה מגמה מירידה לעלייה, וזה קורה בנקודה:

$$A^*=\min\left\{x_i|\sum_{j=1}^{i}p_j\geq\frac{1}{2}\right\}$$

הערך A^* מוגדר להיות החציון.

דוגמה: נניח $X\sim Geo\left(p\right)$ כך ש־$\mathrm{P}\left(X=i\right)=p\left(1-p\right)^{i-1}$.

$$\sum_{j=1}^{i}\mathrm{P}\left(X=j\right)=p+p\left(1-p\right)+...+p\left(1-p\right)^{i-1}=p\frac{1-\left(1-p\right)^i}{1-\left(1-p\right)}=1-\left(1-p\right)^i$$

החציון הוא ערך ה־i המינימלי, המקיים:

$$1-\left(1-p\right)^i\geq\frac{1}{2}$$

או:

$$\frac{1}{2}\geq\left(1-p\right)^i$$

או:

$$\log_{1-p}\left(\frac{1}{2}\right)\leq i$$

כדי לקבל ערך שלם, החציון יהיה:

דוגמא: במ`מ אחיד, החציון יהיה תמיד הערך המרכזי. ואם יש שני ערכים מרכזיים, החציון יהיה כל ערך שביניהם.

20 תיקנון משתנים מקריים

הגדרה: בהינתן מ`מ X, נאמר שהמשתנה המקרי המתוקנן שלו הוא $Z=\frac{X-\mathrm{E}(X)}{\mathrm{SD}(X)}$. כלומר נבצע את הטרנספורמציה הזו על כל התצפיות.
כפי שהגדרנו בסטטיסטיקה תיאורית, המ`מ המתוקנן משמש כדי לתאר את המרחק מהתוחלת ביחידות של סטיית תקן.

דוגמה: נניח כי $X\sim U\left[1,6\right]$, כך ש:

$$\mathrm{E}\left(X\right)=3.5$$

$$\mathrm{Var}\left(X\right)=\frac{35}{12}$$

$$\mathrm{SD}\left(X\right)=\sqrt{\frac{35}{12}}=1.707$$

נחשב את ההתפלגות של המ`מ המתוקנן:

$X=x$	$Z=z$	$\mathrm{P}\left(Z=z\right)$
1	$\frac{-2.5}{1.707}$	$\frac{1}{6}$
2	$\frac{-1.5}{1.707}$	$\frac{1}{6}$
3	$\frac{-0.5}{1.707}$	$\frac{1}{6}$
4	$\frac{0.5}{1.707}$	$\frac{1}{6}$
5	$\frac{1.5}{1.707}$	$\frac{1}{6}$
6	$\frac{2.5}{1.707}$	$\frac{1}{6}$

21 שאלות חזרה

1. בבית מלון קטן שלושה חדרים. שלושת אורחי המלון הגיעו למלון שתויים, וכל אחד לקח את אחד משלושת המפתחות באופן מקרי. נגדיר מ`מ X = מספר האורחים (מבין השלושה) שישנו באותו לילה בחדר הנכון.

(א) מצאו את פונקציית ההסתברות ופונקציית ההתפלגות המצטברת של X. ראשית חשבו אילו ערכים יכול המ`מ X לקבל.

(ב) חשבו את ההסתברויות הבאות:

i. $\mathrm{P}(X\geq 2)$
ii. $\mathrm{P}(X\leq 1)$
iii. $\mathrm{P}(X<1)$

2. בכד ארבעה כדורים שחורים, שלושה אדומים ושלושה לבנים.

(א) מוציאים בזה אחר זה 4 כדורים ללא החזרה. יהי X מס' הכדורים האדומים שהוצאו. מהי פונקציית ההסתברות של X ?

(ב) מוציאים בזה אחר זה כדורים ללא החזרה, עד שמתקבל לראשונה כדור אדום. יהי Y מס' הכדורים הכולל שהוצאו. מהי פונקציית ההסתברות של Y ?

3. במבחן אמריקאי 10 שאלות. לכל שאלה 4 תשובות אפשריות. ראובן לא למד כלל, ומנחש את התשובות באופן מקרי. נגדיר מ`מ X = מס' השאלות עליהן ענה ראובן נכון.

(א) כיצד מתפלג X ? כתבו את פונקצית ההסתברות.

(ב) ציון עובר במבחן זה הוא 55. מהי ההסתברות שראובן יעבור את המבחן?

4. משך החיים של רכיב מסוג מסוים מפולג אחיד על פני 12 חודשים (ההסתברות שהוא יתקלקל לראשונה בכל אחד מהחודשים שווה). במכשיר מותקנים במקביל שלושה מהרכיבים הנ`ל, כאשר המכשיר פועל כל עוד לפחות רכיב אחד מתוכם תקין. משך החיים של כל רכיב בלתי תלוי במשך החיים של הרכיבים האחרים במכשיר.

(א) מהי ההסתברות שהמכשיר יפעל ללא תקלה יותר מחצי שנה?

(ב) 10 מכשירים הופעלו באותו רגע (באופן בלתי תלוי אחד בשני), מהי ההסתברות שכעבור חצי שנה יהיו לפחות 9 מתוכם תקינים?

5. קבוצה של פיראטים תפסה את הבן של המלך כבן ערובה. הם דרשו תיבה של 50 מטבעות זהב בכדי לשחרר את הבן, אחרת יזרקו אותו לכרישים. המלך שלח לפיראטים תיבה והם צריכים לבדוק אם המטבעות הם באמת מזהב. לפיראטים יש זמן לבדוק רק 5 מטבעות מהתיבה לפני שהם חייבים לברוח. הועלו שתי הצעות לבדיקת התיבה:

(א) פיראט א' הציע לבחור 5 מטבעות בזה אחר זה <u>עם החזרה</u> ולבדוק אם הם מזהב, ואם לפחות אחד הוא מזוייף אז הם יזרקו את בן המלך לכרישים.

(ב) פיראט ב' הציע לבחור 5 מטבעות בזה אחר זה <u>ללא החזרה</u> ולבדוק אם הם מזהב, ואם לפחות אחד הוא מזוייף אז הם יזרקו את בן המלך לכרישים.

המלך שלח לפיראטים תיבה עם 40 מטבעות זהב ו־10 מטבעות מזוייפים. הוכיחו כי ההסתברות שהבן יזרק לכרישים היא גבוהה יותר אם יבצעו את הבדיקה שמציע פיראט ב', לעמות ההצעה של פיראט א'. בפרט, חשבו הסתברויות אלו.

6. בידי שיכור ופיכח צרור עם n מפתחות זהים למראה, רק אחד מהם פותח את הדלת לביתם. הפיכח מנסה את המפתחות בזה אחר זה בהסתברות שווה **מתוך אלו שלא ניסה עדיין.** לעומת זאת, בכל ניסיון השיכור בוחר מפתח בהסתברות שווה **ללא קשר למה שכבר ניסה.** נגדיר שני משתנים מקריים X- מספר הנסיונות עד שהפיכח יפתח את הדלת, Y- מספר הניסיונות עד שהשיכור יפתח את הדלת. מה ההתפלגות של X ושל Y? עבור $n=5$ מה ההסתברות שכל אחד מהם יצליח לפתוח את הדלת לאחר שני ניסיונות לכל היותר?

7. מספר הלקוחות המגיעים לבנק בשעה הוא משתנה מקרי פואסון עם פרמטר λ. בבנק יש Kדלפקים ומשך זמן השירות של כל לקוח הוא בדיוק שעה. אם לקוח מגיע ואין דלפק פנוי אז הוא עוזב מיד מבלי לקבל שירות. בתחילת היום כל הדלפקים פנויים. חשבו את פונקציית ההסתברות של שני המשתנים המקריים הבאים:

(א) Y- מספר הלקוחות שהחלו לקבל שירות בשעה הראשונה של היום

(ב) Z- מספר הלקוחות שעזבו ללא קבלת שירות בשעה הראשונה של היום

8. בהגרלת הלוטו באי קטן מגרילים מספר זוכה בודד שנבחר באקראי מתוך $\{1, 2, ..., N\}$ באופן אחיד. בהגרלה משתתפים שני אנשים בלבד, זולו וגולו. כל אחד מהמר על מספר יחיד, והפרס בן S שקלים מחולק באופן שווה בין אלו שניחשו נכונה את המספר הזוכה. זולו וגולו בוחרים את המספר שלהם באקראי (בין 1 ל־N) ובאופן בלתי תלוי.

(א) בכל שבוע יש הגרלה (בשנה יש 52 שבועות). נגדיר X־מספר השבועות בשנה שלפחות אחד מהם (זולו או גולו) זוכה בהגרלה. מהי פונקציית ההסתברות של X?

(ב) מהי תוחלת הרווח של זולו בהגרלה בודדת?

(ג) זולו וגולו רוצים לשפר סיכויים ולכן (לצורך סעיף זה בלבד) הם מחליטים שזולו יבחר במספר 1, גולו יבחר במספר 2 ואם יזכו בפרס כלשהו יחלקו אותו ביניהם. האם השתפרה תוחלת הרווח של כל אחד מהם? הוכיחו.

9. משה רוצה להשתתף בהגרלה של חבילת נופש, כאשר מספר המשתתפים האחרים בהגרלה אינו קבוע אלא הוא מ`מ שמפולג פואסון, $X \sim Poiss(\lambda)$. לכן, הסיכוי לזכות בהגרלה,$Y$, הוא בעצמו משתנה מקרי מפני שהוא פונ' של המ`מ X: $Y = \frac{1}{1+X}$

לשם ההמחשה ניתן לחשוב על כך שההסתברות שמספר המשתתפים חוץ ממשה יהיה 4 היא: $\frac{e^{-\lambda}\lambda^4}{4!}$, לפי פונ' ההסתברות של ההתפלגות הפואסונית. במקרה הזה מתקיים המאורע $\{X = 4\}$, כלומר מספר המשתתפים האחרים הוא 4, והסיכוי לזכייה הוא $\frac{1}{1+x} = \frac{1}{1+4} = \frac{1}{5}$ מכיוון שיש סה`כ 5 אנשים שמשתתפים בהגרלה. מהי תוחלת הסיכוי לזכייה בהגרלה?

10. מושיבים עשר זוגות נשואים סביב שולחן עגול באופן אקראי. מהי תוחלת מספר הנשים שיושבות ליד בעליהן? הדרכה: סמנו את מספר הנשים היושבות ליד בעליהן ב־Y והגדירו את X_i להיות משתנה מקרי שמקבל את הערך 1 אם האישה ה־i יושבת ליד בעלה ו0 אחרת. $i \in \{1, 2...10\}$ מה הקשר בין Y לבין $X_1...X_{10}$? השתמשו בתכונת הלינאריות של התוחלת.

11. נתונים שני מטבעות: מטבע 1 עם הסתברות p_1 ל־`עץ` ומטבע 2 עם הסתברות p_2 ל־`עץ`. נגדיר שני משתנים מקריים:

X־ מספר התוצאות `עץ` כאשר בוחרים מטבע באקראי ומטילים אותו פעמיים

Y־ מספר התוצאות `עץ` כאשר בוחרים מטבע באקראי ומטילים אותו פעם אחת ולאחר מכן בוחרים שוב מטבע באופן אקראי ובלתי תלוי בבחירה הקודמת ומטילים פעם נוספת.

האם התוחלת של X גדולה קטנה או שווה לתוחלת של Y? מה לגבי השונות?

12. נגדיר:

$$\begin{aligned} X &\sim Bin(2, \frac{1}{5}) \\ Y &= X^3 \\ W &= 8 + 3X^3 \end{aligned}$$

חשבו את השונות של Y ושל W

חלק VI
התפלגות משותפת

22 משתנים מקריים רב־ממדיים

בדיון במ`מ אלה, מגדירים לעיתים יותר ממ`מ יחיד.

דוגמה: מטילים מטבע הוגן 3 פעמים. מרחב המדגם מכיל 8 אפשרויות:

$$\Omega=\{(H,H,H),(H,H,T),(H,T,H),(T,H,H),(H,T,T),(T,H,T),(T,T,H),(T,T,T)\}$$

נגדיר את המ`מ X כמספר הראשים, ואת Y כמספר הראשים בהטלה הראשונה פחות מספרם בהטלה השנייה.

	X	Y
H,H,H	3	0
H,H,T	2	0
H,T,H	2	1
T,H,H	2	-1
H,T,T	1	1
T,H,T	1	-1
T,T,H	1	0
T,T,T	0	0

ניתן לבדוק את ההסתברות של X ושל Y בנפרד, ולקבל:

$$\begin{array}{llll} \mathrm{P}(X=0)=\frac{1}{8} & \mathrm{P}(X=1)=\frac{3}{8} & \mathrm{P}(X=2)=\frac{3}{8} & \mathrm{P}(X=3)=\frac{1}{8} \\ \mathrm{P}(Y=-1)=\frac{2}{8} & \mathrm{P}(Y=0)=\frac{4}{8} & \mathrm{P}(Y=1)=\frac{2}{8} & \end{array}$$

אך ניתן גם לבדוק את ההתפלגויות המשותפות:

Y \ X	0	1	2	3
-1	0	$\frac{1}{8}$	$\frac{1}{8}$	0
0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
1	0	$\frac{1}{8}$	$\frac{1}{8}$	0

כך למשל $\mathrm{P}(X=0,Y=-1)=0$ או למשל $\mathrm{P}(X=1.Y=1)=\frac{1}{8}$.

ניתן לחשב את ההתפלגות של X מתוך ההתפלגות המשותפת ל־X ו־Y. בהקשר זה, ההתפלגות של X מכונה `ההתפלגות השולית של $X$`.
נזכור שעבור חלוקה $\{B_j\}_{j=1}^{m}$ מתקיים:

$$\mathrm{P}(X=0)=\sum_{i=1}^{m}\mathrm{P}(X=0,B_i)$$

($(X=0,B_i)$זה חיתוך המאורע $X=0$ ו־ B_i)
ולכן נקבל:

$$\mathrm{P}\left(X=0\right)=\sum_{y}\mathrm{P}\left(X=0,Y=y\right)=0+\frac{1}{8}+0=\frac{1}{8}$$

ובאופן דומה:

$$\mathrm{P}\left(Y=1\right)=\sum_{x}\mathrm{P}\left(X=x,Y=1\right)=0+\frac{1}{8}+\frac{1}{8}+0=\frac{1}{4}$$

וכך ניתן להשלים את כל ההסתברויות השוליות של X ושל Y, ולקבל:

$Y \backslash X$	0	1	2	3	
-1	0	$\frac{1}{8}$	$\frac{1}{8}$	0	$\frac{1}{4}$
0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$
1	0	$\frac{1}{8}$	$\frac{1}{8}$	0	$\frac{1}{4}$
	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	

22.1 פונקציה של משתנים מקריים

בהינתן המשתנים המקריים X,Y, נגדיר את המשתנים המקריים הבאים:

$$Z=X+Y$$

$$W=X-Y$$

$$T=X\cdot Y$$

טענה:

$$\mathrm{E}\left(X\pm Y\right)=\mathrm{E}\left(X\right)\pm\mathrm{E}\left(Y\right)$$

הוכחה: נסמן $p_{xy}=\mathrm{P}\left(X=x,Y=y\right)$, ונזכור, לפי נוסחת ההסתברות השלימה כי $\mathrm{P}(X=x)=\sum_y p_{xy}$ ו־ $\mathrm{P}(Y=y)=\sum_x p_{xy}$ ונחשב:

$$\mathrm{E}\left(X\pm Y\right)=\sum_{x}\sum_{y}\left(x\pm y\right)p_{xy}=\sum_{x}\sum_{y}xp_{xy}\pm\sum_{x}\sum_{y}yp_{xy}=$$

$$=\sum_{x}x\sum_{y}p_{xy}\pm\sum_{y}y\sum_{x}p_{xy}=\sum_{x}x\mathrm{P}\left(X=x\right)\pm\sum_{y}y\mathrm{P}\left(Y=y\right)=$$

$$=\mathrm{E}\left(X\right)\pm\mathrm{E}\left(Y\right)$$

■

הערה: תוצאה זו בעצם אומרת כי $\mathrm{E}\left(X \pm Y\right)$ נקבעת רק בעזרת $\mathrm{E}\left(X\right)$ ו־$\mathrm{E}\left(Y\right)$ ללא חשיבות למידע הנוסף הניתן בהתפלגות המשותפת.
באופן אחר: יתכנו התפלגויות משותפות רבות ל־X ול־Y, אך $\mathrm{E}\left(X \pm Y\right)$ תיקבע רק לפי ההתפלגויות השוליות של X ו־Y.

דוגמה

בכד N כדורים ממוספרים. מוציאים שניים מתוכם, ללא החזרה.
נסמן ב־X את מספרו של הראשון וב־Y את זה של השני.
במקרה $N = 3$ נקבל את ההסתברויות המשותפות הבאות:

X \ Y	1	2	3
1	0	$\frac{1}{6}$	$\frac{1}{6}$
2	$\frac{1}{6}$	0	$\frac{1}{6}$
3	$\frac{1}{6}$	$\frac{1}{6}$	0

ההסתברות השולית היא:

$$\mathrm{P}\left(X=1\right) = \sum_{j=1}^{3} \mathrm{P}\left(X=1, Y=j\right) = 0 + \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

וכן כל ההסתברויות השוליות הן $\frac{1}{3}$.
נשים לב שבדוגמה זו, לכל $a = 1, 2, 3$ מתקיים $\mathrm{P}\left(X=a\right) = \mathrm{P}\left(Y=a\right) = \frac{1}{3}$.

הגדרה: משתנים מקריים X, Y ייקראו **שווי־התפלגות** אם לכל a מתקיים:

$$\mathrm{P}\left(X=a\right) = \mathrm{P}\left(Y=a\right)$$

חשוב להפריד בין `שווי־התפלגות` לבין `שווים`. כלומר, שוויון בהתפלגות <u>לא</u> אומר $\mathrm{P}\left(X=Y\right) = 1$, כפי שראינו בדוגמה האחרונה שבה למעשה $\mathrm{P}\left(X=Y\right) = 0$.

בדוגמה שראינו מתקיים:

$$\mathrm{E}\left(X\right) = \frac{1}{3}\left(1+2+3\right) = 2$$

וכן:

$$\mathrm{Var}\left(X\right) = \left(1-2\right)^2 \frac{1}{3} + \left(2-2\right)^2 \frac{1}{3} + \left(3-2\right)^2 \frac{1}{3} = \frac{2}{3}$$

מכיוון שהתוחלת והשונות הן פונקציה של ההתפלגות, נסיק שהתוחלת והשונות של Y שוות לאלו של X, שכן הם שווי־התפלגות בפרט, $X, Y \sim U(1,3)$.

נמשיך את הדיון בדוגמה: נגדיר $Z = X \cdot Y$. כלומר לכל $\omega \in \Omega$ מתקיים $Z\left(\omega\right) = X\left(\omega\right) \cdot Y\left(\omega\right)$.

$z=xy$	$\mathrm{P}(Z=z)$
1	0
2	$\frac{1}{3}$
3	$\frac{1}{3}$
4	0
6	$\frac{1}{3}$
9	0

נחשב את התוחלת:

$$\mathrm{E}(Z)=\frac{2+3+6}{3}=\frac{11}{3}$$

אך נשים לב:

$$\mathrm{E}(X)\cdot\mathrm{E}(Y)=\left(\frac{2}{3}\right)^2=\frac{4}{9}\neq\frac{11}{3}$$

נסיק מכך שהתוחלת של פונקציה של מ`מ לא בהכרח שווה לפונקציה של התוחלות.
לעומת זאת כפי שראינו לעיל, אם מדובר בפונקציה לינארית התוחלת של הפונקציה שווה לפונקציה של התוחלות:

$t=x+y$	$\mathrm{P}(T=t)$
2	0
3	$\frac{1}{3}$
4	$\frac{1}{3}$
5	$\frac{1}{3}$
6	0

ובחישוב פשוט נקבל עבור המ`מ $T=X+Y$:

$$\mathrm{E}(X)+\mathrm{E}(Y)=2+2=4=\mathrm{E}(X+Y)=\mathrm{E}(T)$$

ניתן לראות ישירות מכך ש-$T\sim U[3,5]$ שהשונות היא כמו השונות של $T-2\sim U[1,3]$:

$$\mathrm{Var}(T)=\frac{3^2-1}{12}=\frac{2}{3}\neq\frac{4}{3}=Var(X)+Var(Y)$$

מסקנה: כאשר הפונקציה לינארית, תוחלת הפונקציה שווה לפונקציה של התוחלת.
כמו־כן שונות הפונקציה לא שווה לפונקציה של השונות, גם כאשר הפונקציה לינארית.

הערה: נגדיר W להיות הערך על הכדור שנותר מבין השלושה. מתקיים $W\sim U[1,3]$.
נשים לב שכל אחד מהבאים X,Y,W הוא מקרי, אולם הסכום שלהם קבוע $X+Y+Z=6$ או:

$$X+Y=6-W$$

נסיק מכך:

$$\mathrm{E}\left(X+Y\right)=\mathrm{E}\left(6-W\right)=\mathrm{E}\left(6\right)-\mathrm{E}\left(W\right)=6-2=4$$

$$\mathrm{Var}\left(X+Y\right)=\mathrm{Var}\left(6-W\right)=\mathrm{Var}\left(-W\right)=\left(-1\right)^{2}\mathrm{Var}\left(W\right)=\mathrm{Var}\left(W\right)=\frac{2}{3}$$

ראינו כי אם $X\sim Bin\left(n,p\right)$ אז $\mathrm{E}\left(X\right)=np$. נראה הוכחה נוספת לכך.
נתבונן באוכלוסייה של n אנשים, ובודקים מי מהם מגיע לבנק. אם הפרט ה־i מגיע לבנק, נגדיר זאת כ`הצלחה`, ונסמנה ב־1.
נגדיר זאת באמצעות `פונקציית אינדיקטור`:

$$1\leq i\leq n\ \ I_{i}=\begin{cases}1 & \text{success in the i-person}\\ 0 & \text{otherwise}\end{cases}$$

נראה כי לכל i:

$$\mathrm{P}\left(I_{i}=1\right)=p$$

$$\mathrm{P}\left(I_{i}=0\right)=1-p$$

$$\mathrm{E}\left(I_{i}\right)=p$$

בפרט, $I_{i}\sim Ber(p)$ $1\leq i\leq n$ והם שווי התפלגות.
נסמן את המ`מ X כמספר ההצלחות, ולכן:

$$X=I_{1}+I_{2}+...+I_{n}=\sum_{i=1}^{n}I_{i}$$

$$\mathrm{E}\left(X\right)=\mathrm{E}\left(\sum_{i=1}^{n}I_{i}\right)=\sum_{i=1}^{n}\mathrm{E}\left(I_{i}\right)=\sum_{i=1}^{n}p=np$$

השוויון השני נובע מכך שתוחלת של סכום היא סכום התוחלות. ■
הערה: נשים לב כי קיבלנו תוחלת של np ללא שימוש בהנחת האי־תלות בין המופעים השונים, כפי שנדרש בהגדרת ההתפלגות בינומית.

דוגמה: התפלגות היפר־גאומטרית

באוכלוסייה קיימים a פרטים מסוג A ו־b פרטים מסוג B. בוחרים מתוכן n פרטים ללא החזרה. $1\leq n\leq a+b$.
נדון במקרה $n=2$ ונגדיר:

$$I_{1}=\begin{cases}1 & \text{if the first is }A\\ 0 & \text{otherwise}\end{cases}$$

$$I_{2}=\begin{cases}1 & \text{if the second is }A\\ 0 & \text{otherwise}\end{cases}$$

נחשב לדוגמה:

$$P\left(I_1=0,I_2=0\right)=P\left(I_1=0\right)P\left(I_2=0|I_1=0\right)=\frac{b}{a+b}\cdot\frac{b-1}{a+b-1}$$

ובאופן דומה נשלים את טבלת ההתפלגות המשותפת כולה:

I_2 \ I_1	0	1	
0	$\frac{b}{a+b}\cdot\frac{b-1}{a+b-1}$	$\frac{a}{a+b}\cdot\frac{b}{a+b-1}$	$\frac{b}{a+b}$
1	$\frac{b}{a+b}\cdot\frac{a}{a+b-1}$	$\frac{a}{a+b}\cdot\frac{a-1}{a+b-1}$	$\frac{a}{a+b}$
	$\frac{b}{a+b}$	$\frac{a}{a+b}$	1

טענה: עבור $X\sim HG\left(n,a,b\right)$ (כלומר אם דוגמים n פרטים ללא החזרה מתוך אוכלוסייה בה יש a מסוג A ו־b מסוג B, והמ`מ X סופר את מספר הנדגמים מסוג A) מתקיים:

$$E\left(X\right)=n\cdot\frac{a}{a+b}$$

הוכחה: עבור פונקציית ההתפלגות:

$$1\le i\le n\,,\ I_i=\begin{cases}1 & \text{if the } i \text{ is } A\\ 0 & \text{otherwise}\end{cases}$$

מתקיים שכל I_i,I_j הם שווי־התפלגות, ובפרט גם:

$$E\left(I_i\right)=E\left(I_j\right)=1\cdot P\left(I_j\in A\right)+0\cdot P\left(I_j\notin A\right)=\frac{a}{a+b}$$

כמו־כן מתקיים לפי הגדרת המ`מ:

$$X=\sum_{i=1}^{n}I_i$$

ולכן נקבל לפי הגדרת התוחלת:

$$E\left(X\right)=E\left(\sum_{i=1}^{n}I_i\right)=\sum_{i=1}^{n}E\left(I_i\right)=\sum_{i=1}^{n}\frac{a}{a+b}=n\cdot\frac{a}{a+b}$$

■

נשים לב שמתקיים לפי השונות של מ`מ ברנולי שהראינו לעיל (עמוד 103):

$$\mathrm{Var}\left(I_i\right)=Var\left(I_j\right)=\frac{a}{a+b}\left(1-\frac{a}{a+b}\right)=\frac{ab}{\left(a+b\right)^2}$$

וכן:

$$E\left(I_i\cdot I_j\right)=P\left(I_i\cdot I_j=1\right)=\frac{a\left(a-1\right)}{\left(a+b\right)\left(a+b-1\right)}$$

דוגמה: נניח כי $X\sim NB\left(r,p\right)$. כלומר מבצעים ניסוי עם הסתברות p להצלחה, עד וכולל ההצלחה ה-r, ו-X סופר את מספר הניסויים. נבדוק מתי תהיה ההצלחה ה-r, מבלי שמשנה לנו התפזרות ההצלחות בדרך. כזכור ראינו שההתפלגות הבינומית השלילית היא:

$$k\geq r\ ,\mathrm{P}\left(X=k\right)=\left(\begin{array}{c} k-1\\ r-1 \end{array}\right)p^{r}\left(1-p\right)^{k-r}$$

טענה:

$$\mathrm{E}\left(X\right)=\frac{r}{p}$$

הוכחה: נגדיר:

X_{1} הוא מספר הניסויים עד ההצלחה הראשונה.

X_{2} הוא מספר הניסויים הנוספים עד ההצלחה הבאה.

$\vdots$

X_{r} הוא מספר הניסויים עד ההצלחה ה-r.

נשים לב כי כולם שווי־התפלגות המקיימים $X_{1},X_{2},...,X_{r}\sim Geo\left(P\right)$, $1\leq i\leq r$, ולכן $\mathrm{E}\left(X_{i}\right)=\frac{1}{p}$.

לפי ההגדרה של המ`מ מתקיים $X=X_{1}+X_{2}+...+X_{r}$, ולכן:

$$\mathrm{E}\left(X\right)=\mathrm{E}\left(X_{1}\right)+\mathrm{E}\left(X_{2}\right)+...+\mathrm{E}\left(X_{r}\right)=\frac{r}{p}$$

$\blacksquare$

דוגמה: חברת קוקה־קולה מציעה מבצע: בכל פקק של בקבוק מופיעה אות עברית מקרית, בהסתברות $\frac{1}{22}$ (האותיות הסופיות ייחשבו כרגילות). כל אדם שמצליח להרכיב את שמו זכאי להשתתף בהגרלת פרס.

נגדיר את מספר הפקקים שיש להשיג כדי להרכיב שם ולהשתתף בהגרלה כמ`מ.

כמה בקבוקים בתוחלת על `איתן` לרכוש כדי להשתתף בהגרלה?

נסמן:

X_{1} מספר הבקבוקים שיש לרכוש עד לאות ראשונה מתאימה.

X_{2} הוא מספר הבקבוקים הנוספים שיש לרכוש עד לאותה השנייה המתאימה.

X_{3} הוא מספר הבקבוקים הנוספים שיש לרכוש עד לאותה השלישית המתאימה.

X_{4} הוא מספר הבקבוקים הנוספים שיש לרכוש עד לאותה האחרונה המתאימה.

נסמן $Y=X_{1}+X_{2}+X_{3}+X_{4}$, ונשים לב ש-$Y$ מסמן את המ`מ שהגדרנו לעיל.

כעת נשים לב להתפלגויות של X_{i}:

$$X_{1}\sim Geo\left(\tfrac{4}{22}\right)\Rightarrow\mathrm{E}\left(X_{1}\right)=\tfrac{22}{4}$$

$$X_{2}\sim Geo\left(\tfrac{3}{22}\right)\Rightarrow\mathrm{E}\left(X_{1}\right)=\tfrac{22}{3}$$

$$X_{1}\sim Geo\left(\tfrac{2}{22}\right)\Rightarrow\mathrm{E}\left(X_{1}\right)=\tfrac{22}{2}$$

$$X_{1}\sim Geo\left(\tfrac{1}{22}\right)\Rightarrow\mathrm{E}\left(X_{1}\right)=\tfrac{22}{1}$$

נסיק מכך:

$$\mathrm{E}(Y)=\mathrm{E}(X_1)+\mathrm{E}(X_2)+\mathrm{E}(X_3)+\mathrm{E}(X_4)=22\left(\tfrac{1}{4}+\tfrac{1}{3}+\tfrac{1}{2}+1\right)$$

ההכללה מכאן לבעלי שמות ארוכים יותר ברורה.

23 קשרים בין משתנים מקריים

23.1 שונות משותפת של משתנים מקריים

הגדרה: בהינתן שני מ`מ X,Y נגדיר ונסמן את השונות המשותפת שלהם:

$$\boxed{\mathrm{Cov}(X,Y)=\mathrm{E}\left[(X-\mathrm{E}(X))(Y-\mathrm{E}(Y))\right]}$$

נשים לב שמההגדרה נובע כי $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$.

נדון בדוגמה בעמוד **117** לכל זוג x,y נבדוק את הערך של $(x-\mathrm{E}(X))(y-\mathrm{E}(Y))$, ונזכור כי $\mathrm{E}(X)=\mathrm{E}(2)=2$:

$X \backslash Y$	1	2	3
1	$(1-2)(1-2)=1$	$(1-2)(2-2)=0$	$(1-2)(3-2)=-1$
2	$(2-2)(1-2)=0$	$(2-2)(2-2)=0$	$(2-2)(3-2)=0$
3	$(3-2)(1-2)=-1$	$(3-2)(2-2)=0$	$(3-2)(3-2)=1$

כעת נוכל לחשב את התוחלת של ההתפלגות שקיבלנו כדי למצוא את השונות המשותפת:

$$\mathrm{Cov}(X,Y)=\mathrm{E}\left[(X-\mathrm{E}(X))(Y-\mathrm{E}(Y))\right]$$

$$=\sum_{x,y}\left[(x-2)(y-2)\right]\cdot \mathrm{P}(X-2)(Y-2)$$

$$=1\cdot 0+0\cdot\tfrac{1}{6}-1\cdot\tfrac{1}{6}+0\cdot\tfrac{1}{6}+0\cdot 0+0\cdot\tfrac{1}{6}-1\cdot\tfrac{1}{6}+0\cdot\tfrac{1}{6}+1\cdot 0=-\tfrac{1}{3}$$

נוסחה:

$$\boxed{\mathbf{Cov}(X,Y)=\mathbf{E}(XY)-\mathbf{E}(X)\,E(Y)}$$

הוכחה:

$$\begin{aligned}\mathrm{Cov}(X,Y) &= \mathrm{E}\left[(X-\mathrm{E}(X))(Y-\mathrm{E}(Y))\right] \\ &= \mathrm{E}\left[XY-\mathrm{E}(Y)X-\mathrm{E}(X)Y+\mathrm{E}(X)\mathrm{E}(Y)\right] \\ &= \mathrm{E}(XY)-\mathrm{E}\left[\mathrm{E}(Y)X\right]-\mathrm{E}\left[\mathrm{E}(X)Y\right]+\mathrm{E}\left[\mathrm{E}(X)\mathrm{E}(Y)\right] \\ &= \mathrm{E}(XY)-\mathrm{E}(Y)\mathrm{E}(X)-\mathrm{E}(X)\mathrm{E}(Y)+\mathrm{E}(X)\mathrm{E}(Y) \\ &= \mathrm{E}(XY)-\mathrm{E}(Y)\mathrm{E}(X)\end{aligned}$$

■

בדוגמה: מתקיים $\mathrm{E}(X)=\mathrm{E}(Y)=2$, וכן חישבנו לעיל $\mathrm{E}(XY)=\frac{11}{3}$, ולכן:

$$\mathrm{Cov}(X,Y)=\frac{11}{3}-4=-\frac{1}{3}$$

תכונות השונות המשותפת:

1. יחידות המדידה של השונות המשותפת הן מכפלת יחידות המדידה של שני המשתנים המקריים.
2. $\mathrm{Cov}(aX+b,cY+d)=ac\cdot\mathrm{Cov}(X,Y)$
3. $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$
4. $\mathrm{Cov}(X,X)=\mathrm{Var}(X)$

דוגמה: באוכלוסייה a פרטים מסוג A, ו-b פרטים מסוג B. בוחרים שני פרטים ללא החזרה. פונקציית ההתפלגות היא:

$$I_1=\begin{cases}1 & \text{if the first is } A\\ 0 & \text{otherwise}\end{cases}$$

$$I_2=\begin{cases}1 & \text{if the second is } A\\ 0 & \text{otherwise}\end{cases}$$

חישבנו את התוחלת:

$$\mathrm{E}(I_1)=\mathrm{E}(I_2)=\frac{a}{a+b}$$

נבדוק מהי השונות המשותפת של שני מ`מ אלו.
אינטואיטיבית ניתן לקבוע שהשונות המשותפת שלילית, שכן אם ידוע כי $I_1=1$ אז

ההסתברות לאירוע $I_2=1$ קטנה יותר. ואכן כפי שנראה זה המצב.
חישוב עזר:

$$I_1\cdot I_2=\begin{cases}1 & \frac{a}{a+b}\cdot\frac{a-1}{a+b-1}\\ 0 & 1-\frac{a}{a+b}\cdot\frac{a-1}{a+b-1}\end{cases}$$

ולכן:

$$\mathrm{E}\left(I_1\cdot I_2\right)=\frac{a}{a+b}\cdot\frac{a-1}{a+b-1}$$

כפי שראינו בעמוד 120.
כעת נחשב את השונות המשותפת:

$$\mathrm{Cov}\left(I_1,I_2\right)=\mathrm{E}\left(I_1\cdot I_2\right)-\mathrm{E}\left(I_1\right)\mathrm{E}\left(I_2\right)=\frac{a}{a+b}\cdot\frac{a-1}{a+b-1}-\left(\frac{a}{a+b}\right)^2$$

$$=\frac{a(a-1)(a+b)-a^2(a+b-1)}{(a+b)^2(a+b-1)}=\frac{a\left(a^2-a+ab-b-a^2-ab+a\right)}{(a+b)^2(a+b-1)}=\frac{-ab}{(a+b)^2(a+b-1)}<0$$

23.2 מקדם המתאם בין משתנים מקריים

הגדרה: בהינתן המ`מ X,Y, מקדם המתאם ביניהם מוגדר ומסומן:

$$\boxed{\mathrm{Corr}\left(X,Y\right)=\frac{\mathrm{Cov}\left(X,Y\right)}{\mathrm{SD}\left(X\right)\cdot\mathrm{SD}\left(Y\right)}}$$

תכונות מקדם המתאם:

1. מקדם המתאם הוא ערך מספרי ללא יחידות מדידה.
2. $$\mathrm{Corr}\left(aX+b,cY+d\right)=\begin{cases}\mathrm{Corr}\left(X,Y\right) & ac>0\\ -\mathrm{Corr}\left(X,Y\right) & ac<0\end{cases}$$

טענה: לכל מ`מ X,Y מתקיים:

$$-1\leq\mathrm{Corr}\left(X,Y\right)\leq 1$$

נוכיח טענה זו בהמשך. (עמוד 126) .

נמשיך לדון בדוגמה הקודמת: נשים לב שמתקיים:

$$I_1, I_2 \sim Bin\left(1, \tfrac{a}{a+b}\right) \Rightarrow \mathrm{Var}\left(I_1\right) = \mathrm{Var}\left(I_2\right) = \tfrac{a}{a+b} \cdot \tfrac{b}{a+b}$$

מכאן:

$$\mathrm{Corr}\left(X, Y\right) = \frac{\mathrm{Cov}\left(I_1, I_2\right)}{\mathrm{SD}\left(I_1\right) \cdot \mathrm{SD}\left(I_2\right)} = \frac{\frac{-ab}{(a+b)^2(a+b-1)}}{\frac{ab}{(a+b)^2}} = \frac{-1}{a+b-1}$$

נשים לב שנובע מכאן כי ככל ש-a, b גדולים יותר, מקדם המתאם (בערך מוחלט) בין I_1, I_2 חלש יותר. ואכן באופן אינטואיטיבי ככל שהאוכלוסיות גדולות יותר, ההשפעה של שליפת פריט אחד על שליפת השני - קטנה יותר.

טרמינולוגיה

X, Y ייקראו **מתואמים שלילית** אם: $\mathrm{Corr}\left(X, Y\right) < 0$.
X, Y ייקראו **מתואמים חיובית** אם: $\mathrm{Corr}\left(X, Y\right) > 0$.
X, Y ייקראו **בלתי־מתואמים** אם:

$$\mathrm{E}\left(XY\right) = \mathrm{E}\left(X\right)\mathrm{E}\left(Y\right) \Leftrightarrow \mathrm{Cov}\left(X, Y\right) = 0 \Leftrightarrow \mathrm{Corr}\left(X, Y\right) = 0$$

.

טענה:

$$\boxed{\mathbf{Var}\left(X \pm Y\right) = \mathbf{Var}\left(X\right) + \mathbf{Var}\left(Y\right) \pm 2\mathbf{Cov}\left(X, Y\right)}$$

הוכחה:

$$\mathrm{Var}\left(X \pm Y\right) = \mathrm{E}\left[X \pm Y - \mathrm{E}\left(X \pm Y\right)\right]^2 = \mathrm{E}\left[X - \mathrm{E}\left(X\right) \pm \left(Y - \mathrm{E}\left(Y\right)\right)\right]^2$$

$$= \mathrm{E}\left(X - \mathrm{E}\left(X\right)\right)^2 \pm 2\mathrm{E}\left(X - \mathrm{E}\left(X\right)\right)\left(Y - \mathrm{E}\left(Y\right)\right) + \mathrm{E}\left(Y - \mathrm{E}\left(Y\right)\right)^2$$

$$= \mathrm{Var}\left(X\right) \pm 2\mathrm{Cov}\left(X, Y\right) + \mathrm{Var}\left(Y\right)$$

■

מסקנה: X, Y בלתי־מתואמים אם ורק אם מתקיים:

$$\mathrm{Var}\left(X \pm Y\right) = \mathrm{Var}\left(X\right) \pm \mathrm{Var}\left(Y\right)$$

מסקנה:

$$\mathrm{Var}\left(aX \pm bY \pm c\right) = \mathrm{Var}\left(aX \pm bY\right) = a^2\mathrm{Var}\left(X\right) + b^2\mathrm{Var}\left(Y\right) \pm 2ab\mathrm{Cov}\left(X,Y\right)$$

מסקנה:

$$\mathrm{Var}\left(X+Y+Z\right)$$

$$= \mathrm{Var}\left(X\right) + \mathrm{Var}\left(Y\right) + \mathrm{Var}\left(Z\right) + 2\mathrm{Cov}\left(X,Y\right) + 2\mathrm{Cov}\left(X,Z\right) + 2\mathrm{Cov}\left(Y,Z\right)$$

נשאיר טענה זו ללא הוכחה.
חישבו על ההכללה של טענה זו לסכום כללי של n משתנים מקריים.

טענה:

$$-1 \le \mathrm{Corr}\left(X,Y\right) \le 1$$

הוכחה: נתבונן בביטוי הבא כפונקציה של t:

$$\mathrm{Var}\left(Y-tX\right) = \mathrm{Var}\left(Y\right) - 2t\mathrm{Cov}\left(X,Y\right) + t^2\mathrm{Var}\left(X\right)$$

זו פונקציה ריבועית ב־t, והיא אי־שלילית.
כזכור עבור פונקציה ריבועית כללית $y = at^2 + bt + c$ אם $a > 0$ היא שלילית לכל ערך של t אם ורק אם $b^2 \le 4ac$ ובמקרה שלנו, אם ורק אם $\mathrm{Cov}^2\left(X,Y\right) \le \mathrm{Var}\left(X\right)\cdot\mathrm{Var}\left(Y\right)$ זאת אומרת, אם ורק אם $-1 \le \mathrm{Corr}\left(X,Y\right) \le 1$.

ערכו של t שממזער את $\mathrm{Var}\left(Y-tX\right)$ הוא:

$$\frac{\mathrm{Cov}\left(X,Y\right)}{\mathrm{Var}\left(X\right)} = \mathrm{Corr}\left(X,Y\right)\cdot\frac{\mathrm{SD}\left(Y\right)}{\mathrm{SD}\left(X\right)}$$

דוגמה: בכד נמצאים $N=3$ כדורים ממוספרים $1,2,3$. מוציאים שניים ללא החזרה.
נגדיר את X כמספר על הכדור הראשון ואת Y כמספר שעל השני.
נחשב את מקדם המתאם.
X,Y שווי־התפלגות ולכן:

$$\mathrm{E}\left(X\right) = \mathrm{E}\left(Y\right) = 2$$

$$\mathrm{Var}\left(X\right) = \mathrm{Var}\left(Y\right) = \tfrac{3^2-1}{12} = \tfrac{2}{3}$$

$$\mathrm{Cov}\left(X,Y\right) = \mathrm{E}\left(XY\right) - \mathrm{E}\left(X\right)\mathrm{E}\left(Y\right) = \tfrac{11}{3} - 2\cdot 2 = -\tfrac{1}{3}$$

נסיק מכאן:

$$\mathrm{Corr}\left(X,Y\right) = \frac{-\frac{1}{3}}{\sqrt{\frac{2}{3}}\cdot\sqrt{\frac{2}{3}}} = -\frac{1}{2}$$

הכללה: נכליל את התוצאה שראינו בדוגמה.
נניח שבכד N כדורים ממוספרים $1,2,...,N$, ומוציאים שניים ללא החזרה.
נגדיר את X_1 כמספר שעל הכדור הראשון ואת X_2 כמספר שעל השני.
נוכיח שמתקיים:

$$\operatorname{Corr}(X_1,X_2)=-\frac{1}{N-1}$$

הוכחה: ברור כי X_1,X_2 שווי־התפלגות, ולכן:

$$\operatorname{E}(X_1)=\operatorname{E}(X_2)=\tfrac{N+1}{2}$$

$$\operatorname{Var}(X_1)=\operatorname{Var}(X_2)=\tfrac{N^2-1}{12}$$

נניח שכל הכדורים הוצאו בזה אחר זה, ונגדיר בהתאם ובנוסף את המ`מ $X_3,...,X_N$.
נטען שלכל זוג מ`מ שנבחר (X_i,X_j), $1\le i\neq j\le N$ התפלגות משותפת זהה עם כל זוג משתנים שנבחר (X_r,X_k), $1\le r\neq k\le N$.
כלומר, למשל ההתפלגות המשותפת של (X_2,X_7) זהה להתפלגות המשותפת של (X_4,X_6) או של (X_2,X_6).
לכן נוכל לסמן באופן כללי את הקבוע:

$$\operatorname{Cov}(X_i,X_j)=C,\ \ i\neq j$$

נשים לב שמתקיים $Var(X_1+X_2+...+X_N)=0$, כי הסכום $X_1+X_2+...+X_N$ הוא ערך קבוע (ושווה ל־$\frac{N(N+1)}{2}$ לפי נוסחת סכום של טור חשבוני), ושונות של קבוע היא 0.
מאידך נזכור:

$$\operatorname{Var}(X+Y+Z)$$

$$=\operatorname{Var}(X)+\operatorname{Var}(Y)+\operatorname{Var}(Z)+2\operatorname{Cov}(X,Y)+2\operatorname{Cov}(X,Z)+2\operatorname{Cov}(Y,Z)$$

וניתן להכליל זאת לכל מספר כלשהו של מ`מ, ולכן:

$$0=\operatorname{Var}(X_1+X_2+...+X_N)=N\operatorname{Var}(X_1)+2\cdot\tfrac{N(N+1)}{2}\cdot C$$

$$=N\tfrac{(N^2-1)}{12}+N(N-1)C$$

וכעת אם נחלץ מהמשוואה את $C=\operatorname{Cov}(X_i,X_j)$ נקבל:

$$C=\frac{-N\left(N^2-1\right)}{12N(N-1)}=\frac{-(N+1)(N-1)}{12(N-1)}=\frac{-(N+1)}{12}$$

ולכן לפי הגדרת מקדם המתאם:

$$\operatorname{Corr}(X,Y)=\frac{\frac{-(N+1)}{12}}{\sqrt{\frac{(N^2-1)}{12}}\sqrt{\frac{(N^2-1)}{12}}}=-\frac{1}{N-1}$$

כלומר ככל ש־N גדול יותר ההשפעה של מ`מ אחד על האחר קטנה יותר. ■

הערה: מתוצאה זו נובע שעבור $N=2$ מתקיים $\mathrm{Corr}\left(X_1,X_2\right)=-1$. כלומר מתקבל מתאם מקסימלי (בערך מוחלט). ההסבר הוא שקיים קשר לינארי בין X_1 ל־X_2 שניתן להציגו $X_2=3-X_1$.

טענה: (ללא הוכחה, נסו בעצמכם)

$$\boxed{\textbf{יהיו } X_1 \textbf{ ו־} X_2 \textbf{ משתנים מקריים ב`ת ש`ה אזי } \mathbf{Var}\left(X\right)=\frac{1}{2}\mathbf{E}\left(X_1-X_2\right)^2}$$

23.3 ישר הרגרסיה בין משתנים מקריים

נבחר a,b, ולאחר שנבדוק את ערכו של X ננבא את ערכו של Y להיות $\hat{Y}=b+aX$.
נשים לב שבחירת a,b היא עוד לפני שידענו את ערכו של X.
תוחלת ריבוע השגיאה בין הערך המנובא לערך האמתי היא:

$$\mathrm{E}\left[Y-\left(b+aX\right)\right]^2$$

כעת נחפש מקדמים a,b שימזערו תוחלת ריבוע השגיאה.

טענה:

$$a_{\min}=\mathrm{Corr}\left(X,Y\right)\cdot\frac{\mathrm{SD}(X)}{\mathrm{SD}(Y)}$$

$$b_{\min}=\mathrm{E}\left(Y\right)-a_{\min}\mathrm{E}\left(X\right)$$

הוכחה: נזכור כי:

$$\mathrm{Var}\left(X\right)=\mathrm{E}\left(X^2\right)-\mathrm{E}^2\left(X\right)\Rightarrow\mathrm{E}\left(X^2\right)=\mathrm{Var}\left(X\right)+E^2\left(X\right)$$

ומכאן שמתקיים:

$$\mathrm{E}\left(Y-b-aX\right)^2=\mathrm{Var}\left(Y-b-aX\right)+\mathrm{E}^2\left(Y-b-aX\right)$$

$$=\mathrm{Var}\left(Y-aX\right)+\mathrm{E}^2\left(Y-b-aX\right)$$

ראשית נשים לב כי המחובר הראשון אינו פונקציה של b. יתר על כן כאשר $b=\mathrm{E}\left(Y\right)-a\mathrm{E}\left(X\right)$ המחובר השני יתאפס, ולכן זה אידאלי כדי למזער את הביטוי. לבסוף נראה כי $\mathrm{Var}(Y-aX)$ היא פונקציה ריבועית ב־a ונמצא את המינימום שלה:

$$\mathrm{Var}(Y-aX)=\mathrm{Var}(Y)-2a\cdot\mathrm{Cov}(X,Y)+a^2\mathrm{Var}(X)$$

$$a_{\min}=\frac{-2\cdot\mathrm{Cov}(X,Y)}{2\cdot\mathrm{Var}(X)}=\mathrm{Corr}(X,Y)\cdot\frac{\mathrm{SD}(Y)}{\mathrm{SD}(X)}$$

■

אם נציב ערך זה של a נקבל, לאחר קצת אלגברה, כי

$$\mathrm{E}(Y-b-a_{min}X)^2=\mathrm{Var}(Y-a_{min}X)=(1-\mathrm{Corr}^2(X,Y))\mathrm{Var}(Y)$$

הישר $y=b_{\min}+a_{\min}x$ נקרא **ישר הרגרסיה של** Y **על** X.

דוגמה: נמשיך עם הדוגמה הקודמת (עמוד 117) – מוציאים שני כדורים ללא החזרה, מתוך כד של N כדורים ממוספרים. מגדירים את X להיות המספר שעל הכדור הראשון ואת Y להיות המספר שעל השני.
נחשב את ישר הרגרסיה של Y על X:

$$y = b + ax$$

$$a = \mathrm{Corr}\left(X,Y\right) \cdot \tfrac{\mathrm{SD}(X)}{\mathrm{SD}(Y)}$$

$$b = \mathrm{E}\left(Y\right) - a\mathrm{E}\left(X\right)$$

ראינו לעיל שבדוגמה זו מתקיים:

$$\mathrm{Corr}\left(X,Y\right) = \frac{-1}{N-1}$$

ומאחר ש-X,Y שווי־התפלגות אז $\mathrm{SD}\left(X\right) = \mathrm{SD}\left(Y\right)$ ולכן $\tfrac{\mathrm{SD}(X)}{\mathrm{SD}(Y)} = 1$.
נסיק מכך את ערכו של a:

$$a = \frac{-1}{N-1}$$

כמו־כן נזכור שמתקיים:

$$\mathrm{E}\left(X\right) = \mathrm{E}\left(Y\right) = \frac{N+1}{2}$$

ולכן נקבל את ערכו של b:

$$b = \frac{N+1}{2} + \frac{1}{N-1} \cdot \frac{N+1}{2} = \frac{N^2 - 1 + N + 1}{2\left(N-1\right)} = \frac{N\left(N+1\right)}{2\left(N-1\right)}$$

מכאן שישר הרגרסיה הוא:

$$y = \frac{N\left(N+1\right)}{2\left(N-1\right)} - \frac{1}{N-1} \cdot x$$

נשים לב שברגע שיש הרבה כדורים, המידע על תוצאת הראשון לא מספק הרבה מידע על התוצאה של השני. זה ניכר בביטוי שקיבלנו בכך שכאשר $N \to \infty$ אז השיפוע שואף ל-0, והיחס בין החותך, b, לבין התוחלת של Y (השווה ל$\frac{N}{2}$) שואף ל-1.

שונות מוסברת ושונות לא־מוסברת

ראינו שעבור ישר הרגרסיה, הביטוי:

$$\mathrm{E}\left(Y - b - aX\right)^2 = \left(1 - \mathrm{Corr}^2\left(X,Y\right)\right)\mathrm{Var}\left(Y\right)$$

הוא השונות הבלתי־מוסברת של Y, כפונקציה לינארית ב-X. כלומר גודל הטעות בניבוי, במונחי תוחלת.
באופן טבעי, השונות המוסברת תהיה $\mathrm{E}\left(b + aX - \mathrm{E}\left(Y\right)\right)^2$.

טענה:

$$\mathrm{E}\left(b+aX-\mathrm{E}\left(Y\right)\right)^{2}=\mathrm{Corr}^{2}\left(X,Y\right)\cdot\mathrm{Var}\left(Y\right)$$

הוכחה: נוכיח זאת לפי הנוסחה $\mathrm{E}\left(T^{2}\right)=Var\left(T\right)+\mathrm{E}^{2}\left(T\right)$:

$$\mathrm{E}\left(b+aX-\mathrm{E}\left(Y\right)\right)^{2}=\mathrm{Var}\left(b+aX-\mathrm{E}\left(Y\right)\right)+\underbrace{\mathrm{E}^{2}\left(b+aX-\mathrm{E}\left(Y\right)\right)}_{=0}$$

$$\Downarrow$$

$$\mathrm{E}\left(b+aX-\mathrm{E}\left(Y\right)\right)^{2}=\mathrm{Var}\left(b+aX-\mathrm{E}\left(Y\right)\right)$$

כמו־כן מתקיים:

$$\mathrm{Var}\left(b+aX-\mathrm{E}\left(Y\right)\right)=\mathrm{Var}\left(aX\right)=a^{2}\mathrm{Var}\left(X\right)$$

$$=\mathrm{Corr}^{2}\left(X,Y\right)\cdot\frac{Var(Y)}{Var(X)}\cdot\mathrm{Var}\left(X\right)=\mathrm{Corr}^{2}\left(X,Y\right)\cdot\mathrm{Var}\left(Y\right)$$

ומכאן השוויון המבוקש. ■

לסיכום, השונות של Y שווה ל:
{השונות המוסברת של Y על־ידי פונקציה לינארית ב־X}+{השונות הבלתי־מוסברת של Y על־ידי פונקציה לינארית ב־X}.

24 אי־תלות בין משתנים מקריים

הגדרה: המשתנים המקריים $X_1,X_2,...,X_n$ ייקראו **בלתי־תלויים** אם לכל $x_1,x_2,...,x_n$ מתקיים:

$$\mathrm{P}\left(X_1=x_1,X_2=x_2,...,X_n=x_n\right)=\mathrm{P}\left(X_1=x_1\right)\cdot\mathrm{P}\left(X_2=x_2\right)\cdot...\cdot\mathrm{P}\left(X_n=x_n\right)$$

דוגמה:

X_2 \ X_1	0	1	
0	0.12	0.18	0.3
1	0.28	0.42	0.7
	0.4	0.6	

כך למשל ניתן לראות שמתקיים:

$$\mathrm{P}\left(X_1=1,X_2=0\right)=0.18=0.6\cdot0.3=\mathrm{P}\left(X_1=1\right)\mathrm{P}\left(X_2=0\right)$$

טענה: אם X_1, X_2, X_3 בלתי־תלויים, אז גם כל זוג מביניהם בלתי־תלוי.
ובאופן כללי, אם $X_1, X_2, ..., X_n$ מ`מ בלתי־תלויים, אז גם כל קבוצה חלקית שלהם מורכבת ממ`מ בלתי־תלויים.

הוכחה: לכל x_1, x_2 מתקיים:

$$\mathrm{P}\left(X_1 = x_1, X_2 = x_2\right) = \sum_{x_3} \mathrm{P}\left(X_1 = x_1, X_2 = x_2, X_3 = x_3\right)$$

$$\sum_{x_3} \mathrm{P}\left(X_1 = x_1\right) \mathrm{P}\left(X = x_2\right) \mathrm{P}\left(X_3 = x_3\right) = \mathrm{P}\left(X_1 = x_1\right) \mathrm{P}\left(X = x_2\right) \sum_{x_3} \mathrm{P}\left(X_3 = x_3\right)$$

$$= \mathrm{P}\left(X_1 = x_1\right) \mathrm{P}\left(X = x_2\right) \cdot 1 = \mathrm{P}\left(X_1 = x_1\right) \mathrm{P}\left(X = x_2\right)$$

■

טענה: אם X_1, X_2 מ`מ ב`ת אז $\mathrm{Corr}\left(X_1, X_2\right) = 0$.

הוכחה: נוכיח את הטענה השקולה $\mathrm{E}\left(X_1 X_2\right) = \mathrm{E}\left(X_1\right) \mathrm{E}\left(X_2\right)$:

$$\mathrm{E}\left(X_1 X_2\right) = \sum_{x_1} \sum_{x_2} x_1 x_2 \mathrm{P}\left(X_1 = x_1, X_2 = x_2\right)$$

$$= \sum_{x_1} x_1 \mathrm{P}\left(X_2 = x_2\right) \sum_{x_2} x_2 \mathrm{P}\left(X_1 = x_1\right) = \mathrm{E}\left(X_1\right) \mathrm{E}\left(X_2\right)$$

■

הערה: הטענה ההפוכה אינה נכונה. כלומר, אם $Corr\left(X_1, X_2\right) = 0$ זה לא אומר שהם בלתי־תלויים.

דוגמה־נגדית: ניקח סדרה של זוגות נתונים $(-1,0), (0,-1), (1,0), (0,1)$, כאשר בכל זוג הערך הראשון הוא X_1 והשני הוא X_2. כל זוג מתקבל בהסתברות 0.25.
נשים לב שמתקיים:

$$\mathrm{E}\left(X_1\right) = \mathrm{E}\left(X_1\right) = 0$$

$$X_1 \cdot X_2 = 0$$

$$Cov\left(X_1, X_2\right) = 0 - 0 = 0$$

וכעת נחשב את הדוגמה הנגדית:

$$\frac{1}{4} = \mathrm{P}\left(X_1 = -1, X_2 = 0\right) \neq \mathrm{P}\left(X_1 = -1\right) \mathrm{P}\left(X_2 = 0\right) = \frac{1}{4} \cdot \frac{1}{2}$$

מסקנה בעזרת הטענה מעמוד 125 אם X, Y מ`מ ב`ת אזי $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

מסקנה בעזרת עמוד 119 עבור $X \sim Bin(n,p)$ מתקיים $\mathrm{Var}(X) = np(1-p)$

**התפלגות סכום של מ`מ בינומיים:** נניח כי $X_1 \sim Bin\left(n,p\right)$, $X_2 \sim Bin\left(m,p\right)$, והם ב`ת.
אז מתקיים כי $X_1 + X_2 \sim Bin\left(n+m,p\right)$.
הוכחה: כדי להוכיח שמ`מ מתפלג באופן כלשהו, צריך להראות שפונקצייה ההתפלגות שלו מתאימה להגדרת ההתפלגות.
כלומר במקרה זה צריך להראות שלכל $0 \le i \le n+m$ מתקיים:

$$\mathrm{P}\left(X_1 + X_2 = i\right) = \left(\begin{array}{c} n+m \\ i \end{array}\right) p^i \left(1-p\right)^{n+m-i}$$

נחשב:

$$\mathrm{P}\left(X_1 + X_2 = i\right) = \sum_{k=0}^{i} \mathrm{P}\left(X_1 = k, X_2 = i-k\right) =$$

$$= \textstyle\sum_{k=0}^{i} \mathrm{P}\left(X_1 = k\right) \mathrm{P}\left(X_2 = i-k\right)$$

$$= \sum_{k=0}^{i} \left(\begin{array}{c} n \\ k \end{array}\right) p^k \left(1-p\right)^{n-k} \left(\begin{array}{c} m \\ i-k \end{array}\right) p^{i-k} \left(1-p\right)^{m-i+k}$$

$$= p^i \left(1-p\right)^{m-i+k} \textstyle\sum_{k=0}^{i} \left(\begin{array}{c} n \\ k \end{array}\right) \left(\begin{array}{c} m \\ i-k \end{array}\right)$$

$$= p^i \left(1-p\right)^{m-i+k} \sum_{k=0}^{i} \left(\begin{array}{c} n \\ k \end{array}\right) \left(\begin{array}{c} m \\ i-k \end{array}\right) \cdot \frac{\left(\begin{array}{c} n+m \\ i \end{array}\right)}{\left(\begin{array}{c} n+m \\ i \end{array}\right)}$$

$$= \left(\begin{array}{c} n+m \\ i \end{array}\right) p^i \left(1-p\right)^{n+m-i}$$

המעבר האחרון נובע מכך שמתקיים $1 = \sum_{k=0}^{i} \frac{\left(\begin{array}{c} n \\ k \end{array}\right)\left(\begin{array}{c} m \\ i-k \end{array}\right)}{\left(\begin{array}{c} n+m \\ i \end{array}\right)}$, שכן זהו סכום ההסתברויות של מ`מ היפר־גאומטרית כפי שהראינו בעמוד 95 . ■
הערה: שימו לב שכאשר $k \ge n$, $i-k \ge m$ או $i-k \le 0$, ערכו של המקדם הקומבינטורי המתאים הוא 0..

הוכחה שנייה: מ`מ בינומי הוא סכום של מ`מ ברנוליים שווי־התפלגות וב`ת. ולכן הוא ניתן להצגה באופן הבא:

$$X_1 = \sum_{i=1}^{n} I_i$$

$$X_2 = \sum_{i=n+1}^{n+m} I_i$$

כאשר מגדירים:

$$I_i = \begin{cases} 1 & p \\ 0 & 1-p \end{cases}$$

ולכן $i \neq j$ המ`מ $I_i, I_j$ ב`ת.
מכאן ניתן להסיק:

$$X_1 + X_2 = \sum_{i=1}^{n+m} I_i$$

ובמילים: סכום של $n+m$ מ`מ ב`ת מפולגים ברנולית עם פרמטר p, כלומר מ`מ מפולג בינומית עם פרמטרים $n+m$ ו–p. ■

טענה: אם X מתפלג בינומי עם פרמטרים n ו־ p אזי:

$$\operatorname{Var}(X) = np(1-p)$$

הוכחה: נחשב:

$$\operatorname{Var}(X) = \operatorname{Var}\left(\sum_{i=1}^{n} I_i\right) = \sum_{i=1}^{n} \operatorname{Var}(I_i) = \sum_{i=1}^{n} p(1-p) = np(1-p)$$

השוויון השני נובע מכך שמדובר במ`מ ב`ת. ■

טענה: אם $X \sim Pois(\lambda)$ אז $\operatorname{Var}(X) = \lambda$.

הוכחה: פואסון הוא קירוב בינומי של $X \sim Bin(n,p)$, ולכן:

$$\operatorname{Var}(X) = np(1-p) = np - np^2 \underset{\substack{n\to\infty \\ p\to 0 \\ np\to\lambda}}{\longrightarrow} \lambda - 0 = \lambda$$

■

25 שקלול אופטימלי בין משתנים מקריים

נניח שמתעניינים בהכנסה הממוצעת של שכירים. ממוצע זה יתקבל אם נתבונן באוכלוסיה כולה ונחשב את הממוצע מתוך נתוני ההכנסות. כך גם נוכל לקבל את שונות ההכנסות. נסמן את ממוצע ההכנסות ב–μ ואת שונותן ב–σ^2. לפני שאנו יודעים מהם μ, σ^2, נדגום מתוך האוכלוסייה נדגם מקרי בודד, נסמן את הכנסתו ב–X_1, אם כל אחד מהפרטים נדגם

בהסתברות שווה מתקיים:

$$\mathrm{E}\left(X_1\right)=\mu$$

$$\operatorname{Var}\left(X_1\right)=\sigma^2$$

נרצה לאמוד את ערכו של μ, ונעשה זאת באמצעות ממוצע שיילקח על־פני מדגם מסוים.
אנו חשים כי ככל שניקח יותר נדגמים ב`ת כך נאמוד בצורה טובה יותר את ערכו של μ.
אם ניקח למשל שני נדגמים ב`ת ונמצע עליהם, נקבל:

$$\mathrm{E}\left(\frac{X_1+X_2}{2}\right)=\frac{1}{2}\left(\mathrm{E}\left(X_1\right)+\mathrm{E}\left(X_2\right)\right)=\frac{1}{2}(\mu+\mu)=\mu$$

$$\operatorname{Var}\left(\frac{X_1+X_2}{2}\right)=\frac{1}{4}\cdot\left(\operatorname{Var}\left(X_1\right)+\operatorname{Var}\left(X_2\right)\right)=\frac{2\sigma^2}{4}=\frac{\sigma^2}{2}$$

ככלל, התוחלת קבועה לכל מספר של נדגמים, אולם השונות תקטן ככל שנגדיל את מספר הנדגמים, כי ערכה הוא $\frac{\sigma^2}{n}$, כפי שנראה מייד.
הקיטון בשונות מעניק משמעות פורמלית לתחושה שככל שהמדגם גדול יותר השגיאה של $\overline{X}$ באמידת μ קטנה יותר.
ואכן מתקיים:

$$\operatorname{Var}(\overline{X})=\mathrm{E}\left(\overline{X}-\mu\right)^2=\operatorname{Var}\left(\overline{X}-\mu\right)+\mathrm{E}^2\left(\overline{X}-\mu\right)=\operatorname{Var}\left(\overline{X}\right)+0$$

$$=\operatorname{Var}\left(\overline{X}\right)=\operatorname{Var}\left(\tfrac{1}{n}\sum_{i=1}^{n}X_i\right)=\tfrac{1}{n^2}\cdot n\cdot\operatorname{Var}\left(X\right)=\tfrac{\sigma^2}{n}$$

במידה ונרצה שלא לתת משקל שווה לכל נדגם, כך שהאומד ל-μ יהיה $\alpha X_1+(1-\alpha)X_2$ עבור α כלשהו,
התוחלת לא תשתנה:

$$\mathrm{E}\left[\alpha X_1+(1-\alpha)X_2\right]=\alpha\mathrm{E}\left(X\right)+(1-\alpha)\mathrm{E}\left(X_2\right)=\mathrm{E}\left(X_1\right)=\mu$$

ואז השאלה היא מה צריך להיות ערכו של α כך ששונות $\alpha X_1+(1-\alpha)X_2$ תהיה מינימלית.
מיד נראה כי הבחירה האופטימלית היא $\alpha=0.5$ אך לפני כן נכניס שאלה זו לבעיה יותר רחבה.
נניח בלי הגבלת הכלליות כי $\operatorname{Var}(X_1)\leq\operatorname{Var}(X_2)$ (אבל עדיין $\mathrm{E}\left(X_1\right)=\mathrm{E}\left(X_2\right)$. ו$X_1$ ו-X_2 ב`ת) מה הבחירה האופטימלית של α כעת? נראה במבט ראשון כי $\alpha=1$ היא התשובה, אחרי הכל, הוא בעל השונות הנמוכה יותר. אבל במבט שני נבחין כי בחירה זו למעשה מתעלמת מאינפורמציה הנתונה ב-X_2 ואולי תשמש אותנו לקבלת שונות נמוכה יותר מאשר $\operatorname{Var}(X_1)$. פורמלית בעייתנו היא למצוא α כך ש $\operatorname{Var}(\alpha X_1+(1-\alpha)X_2)$ מינימלי.
ניתן לראות כי מדובר בפונקציה ריבועית ב-α:

$$\operatorname{Var}\left[\alpha X_1+(1-\alpha)X_2\right]=\alpha^2\operatorname{Var}\left(X_1\right)+(1-\alpha)^2\operatorname{Var}\left(X_2\right)$$

$$=\alpha^2\left[\operatorname{Var}\left(X_1\right)+\operatorname{Var}\left(X_2\right)\right]-2\alpha\operatorname{Var}\left(X_2\right)+\operatorname{Var}\left(X_2\right)$$

והיא מקבלת מינימום ב:

$$\alpha^* = \frac{\text{Var}\,(X_2)}{\text{Var}\,(X_1) + \text{Var}\,(X_2)}$$

$$1 - \alpha^* = \frac{\text{Var}\,(X_1)}{\text{Var}\,(X_1) + \text{Var}\,(X_2)}$$

ניתן לראות כי כל משתנה מקבל משקל פרופורציונלי לשונות השני. דרך אחרת היא לראות כי:

$$\alpha^* = \frac{\frac{1}{\text{Var}(X_1)}}{\frac{1}{\text{Var}(X_1)} + \frac{1}{\text{Var}(X_2)}}$$

$$1 - \alpha^* = \frac{\frac{1}{\text{Var}(X_2)}}{\frac{1}{\text{Var}(X_1)} + \frac{1}{\text{Var}(X_2)}}$$

ז`א כל אחד מקבל משקל שהוא פרופורציונלי להופכי של שונותו. נקודת מבט זו ניתנת להכללה במקרה של מספר רב יותר של משתנים מקריים המעורבים בשקלול
נציב ערך זה של α^* בשונות $\text{Var}\left[\alpha X_1 + (1-\alpha)\, X_2\right]$ ומעט אלגברה תביא אותנו לכך ששונות זו שווה ל:

$$\frac{1}{\frac{1}{\text{Var}(X_1)} + \frac{1}{\text{Var}(X_2)}}$$

קרי למחצית הממוצע ההרמוני בין שונויות X_1 ו-X_2 (ראה עמוד 10),
תוצאה זו ניתנת להרחבה מידית למספר רב של משתנים מקריים, נאמר n. בפרט, כל משתנה מקרי מקבל משקל פרופורציוני הפוך לשונותו והשונות המינימלית . המינימום שווה לממוצע ההרמוני של השונויות מחולק ב־n.

הערה: הבחירה התת אופטימלית $\alpha = 0.5$ מביאה אותנו ל:

$$\text{Var}\left[\frac{1}{2}X_1 + \frac{1}{2}X_2\right] = \frac{1}{4}\text{Var}(X_1) + \frac{1}{4}\text{Var}(X_2) = \frac{1}{2}[\frac{\text{Var}(X_1) + \text{Var}(X_2)}{2}]$$

קרי למחצית הממוצע החשבוני בין השונויות. כך ראינו כבדרך אגב כי הממוצע ההרמוני תמיד קטן או שווה לממוצע החשבוני.

השאלה הופכת למורכבת יותר כאשר X_1 ו-X_2 אינם בלתי תלויים, או יותר נכון לא בלתי מתואמים. נסמן ב-$\text{Cov}(X_1, X_2)$ את השונות המשותפת. אז מטרתנו היא שוב

$$\alpha^* = \underset{\alpha}{\text{argmin}}\text{Var}\left[\alpha X_1 + (1-\alpha)\, X_2\right]$$

טענה:

$$\alpha_{\min} = \frac{\text{Var}\,(X_2) - \text{Cov}\,(X_1, X_2)}{\text{Var}\,(X_1) + \text{Var}\,(X_2) - 2\text{Cov}\,(X_1, X_2)}$$

הוכחה:

$$\mathrm{Var}\left[\alpha X_1 + (1-\alpha) X_2\right] =$$

$$\alpha^2 \mathrm{Var}\left(X_1\right) + (1-\alpha)^2 \mathrm{Var}\left(X_2\right) + 2\alpha (1-\alpha) \mathrm{Cov}\left(X_1, X_2\right) =$$

$$\alpha^2 \mathrm{Var}\left(X_1\right) + \left(1 - 2\alpha + \alpha^2\right) \mathrm{Var}\left(X_2\right) + \left(2\alpha - 2\alpha^2\right) \mathrm{Cov}\left(X_1, X_2\right) =$$

$$\alpha^2 \left(\mathrm{Var}\left(X_1\right) + \mathrm{Var}\left(X_2\right) - 2\mathrm{Cov}\left(X_1, X_2\right)\right) =$$

$$-2\alpha \left(\mathrm{Var}\left(X_2\right) - \mathrm{Cov}\left(X_1, X_2\right)\right) + \mathrm{Var}\left(X_2\right)$$

כפי שראינו השונות של הסכום היא פונקציה ריבועית ב-α, ערכו של $\alpha_{\min}$ נקבע לפי הנוסחה למינימום של פונקציה מסוג זה. ניתן לראות כי בעזרת גזירה $\alpha_{\min}$ הוא פונקציה מונוטנית עולה ב- $\mathrm{Cov}\left(X_1, X_2\right)$ ז`א ככל שמקדם המתאם בין X_1 ל- X_2 גדול יותר כך ניתן משקל גדול יותר ל- X_1 . ■

ניתן לראות גם כי:

$$\mathrm{Var}\left[\alpha X_{\min 1} + \left(1 - \alpha_{\min}\right) X_2\right] = \mathrm{Var}\left(X_2\right) - \alpha_{min}(\mathrm{Var}\left(X_2\right) - \mathrm{Cov}\left(X_1, X_2\right)) =$$

$$\mathrm{Var}\left(X_2\right) - \frac{\left(\mathrm{Var}(X_2) - \mathrm{Cov}(X_1, X_2)\right)^2}{\mathrm{Var}(X_1) + \mathrm{Var}(X_2) - 2\mathrm{Cov}(X_1, X_2)}$$

באופן דומה ניתן לראות כי השונות האופטימלית מונוטונית עולה ב-$\mathrm{Cov}\left(X_1, X_2\right)$. ז`א ככל ש $\mathrm{Cov}(X_1, X_2)$ נמוך יותר (ובפרט, יותר ויותר שלילי) כך השונות האופטימלית נמוכה יותר. לכן, עדיף לשקלל בין אומדים עם מתאם שלילי מאשר אילו עם מתאם חיובי.

26 שאלות חזרה

1. בכד ישנם שלושה כדורים ממוספרים $\{0, 1, 2\}$. מוציאים שני כדורים עם החזרה ורושמים את התוצאות.
נגדיר את שני המשתנים המקריים הבאים: X - המספר הנמוך ביותר שנבחר, Y - מספר הכדור הגבוה ביותר שנבחר.
   (א) רשמו את פונקצית ההסתברות המשותפת של X ו-Y.
   (ב) חשבו את ההתפלגות השולית ואת התוחלות של X ו-Y. האם X ו-Y הם שווי התפלגות?
   (ג) חשבו את התוחלות של הסכום $X+Y$ ושל המכפלה XY.

2. יהיו X, Y, Z משתנים מקריים, ונתונה פונקצית ההסתברות המשותפת:

$$\mathrm{P}(X=1, Y=2, Z=3) = \frac{1}{4}$$

$$\mathrm{P}(X=2, Y=1, Z=1) = \frac{1}{4}$$

$$\mathrm{P}(X=2,Y=2,Z=1)=\frac{1}{4}$$

$$\mathrm{P}(X=2,Y=3,Z=2)=\frac{1}{4}$$

(א) רשמו את התומכים של X,Y,Z.

(ב) חשבו את ההתפלגויות השוליות ואת התוחלות של Z, Y ו־X, האם הם שווי התפלגות?

(ג) חשבו את התוחלות $\mathrm{E}(XYZ)$ ו־ $\mathrm{E}(XY+XZ+YZ)$.

3. מטילים מטבע הוגן 6 פעמים באופן בלתי תלוי. נגדיר X_i־ משתנה מקרי המקבל את הערך 1 אם בהטלה ה־ i התקבל עץ ו־0 אם בהטלה ה־ i התקבל פלי, $i\in\{1,2...6\}$

(א) יהי $S=\sum_{i=1}^{6}X_i$ חשבו את $\mathrm{E}(S)$

(ב) חשבו את $\mathrm{P}(X_1=X_3=X_5=X_6)$

(ג) נגדיר: $Y=X_2\cdot X_4\cdot X_6$ חשבו את פונקצית ההסתברות המשותפת של Y ו־ S. (אפשר לרשום את התשובה כטבלה)

4. מטילים קוביה הוגנת n פעמים. נגדיר את שני המשתנים המקריים הבאים:

X־ מספר הפעמים שהתוצאה היא 1

Y־ מספר הפעמים שהתוצאה היא 6

חשבו את השונות המשותפת ומקדם המתאם של X ו־Y

5. נניח כי X_1,X_2 הם משתנים מקריים בלתי תלויים המתפלגים $\mathrm{Bin}(1,1/2)$. נסמן $X=|X_1-X_2|$ ו־$Y=X_1\cdot X_2$.

(א) סמנו את הטענה הנכונה:

i. X,Y הם בלתי תלויים ולכן גם בלתי מתואמים.

ii. X,Y אינם בלתי תלויים, אך הם בלתי מתואמים.

iii. X,Y אינם בלתי מתואמים ולכן גם אינם בלתי תלויים.

iv. X,Y הם שווי התפלגות.

(ב) סמנו את הטענה הנכונה:

i. X,X_1 הם בלתי תלויים ולכן גם בלתי מתואמים

ii. X,X_1 אינם בלתי תלויים, אך הם בלתי מתואמים.

iii. X,X_1 אינם בלתי מתואמים ולכן גם אינם בלתי תלויים

iv. X,X_1 אינם שווי התפלגות.

6. נתונה פונקצית ההסתברות המשותפת הבאה:

Y \ X	0	1	2	
0	0.03	0.02	0.05	0.1
1	0.09	0.06	0.15	0.3
2	0.18	0.12	0.3	0.6
	0.3	0.2	0.5	

האם X ו־ Y ב`ת? אם כן, הראו שהשוויון: $\mathrm{E}(XY)=\mathrm{E}(X)\mathrm{E}(Y)$ מתקיים

7. משקיע מעוניין לשהקיע סכום כסף מסוים בשתי מניות X_1 ו־ X_2 בעלות תוחלת רווח $\mu_1 = 0.075$ ו־ $\mu_2 = 0.1$, שוניות $\sigma_1^2 = 0.01$ ו־ $\sigma_2^2 = 0.008$ ושונות משותפת: $\mathrm{Cov}(X_1, X_2) = -0.0009$. המשקיע שונא סיכון ולכן מעוניין למזער את שונות תיק ההשקעות שלו. מצא את תיק ההשקעות מהצורה:

$$X = aX_1 + (1-a)X_2$$

אשר ממזער את שונות התשואה על ההשקעה, כלומר :

$$a^* = \underset{a}{\mathrm{argmin}}\mathrm{Var}(aX_1 + (1-a)X_2)$$

מהן תוחלת ושונות התשואה של התיק שנבחר?

חלק VII
אי־שוויונים

27 אי־שוויון מרקוב

משפט: נניח שהמ`מ X יכול לקבל רק ערכים אי־שליליים, אזי לכל $a>0$ מתקיים:

$$\mathrm{P}\left(X\geq a\right)\leq\frac{\mathrm{E}\left(X\right)}{a}$$

ובאופן שקול:

$$\mathrm{P}\left(X<a\right)\geq1-\frac{\mathrm{E}\left(X\right)}{a}$$

הוכחה: נחשב:

$$\mathrm{E}\left(X\right)=\sum_{x}x\cdot\mathrm{P}\left(X=x\right)\geq\sum_{x\geq a}x\cdot\mathrm{P}\left(X=x\right)$$

$$\geq\sum_{x\geq a}a\cdot\mathrm{P}\left(X=x\right)=a\cdot\sum_{x\geq a}\mathrm{P}\left(X=x\right)=$$

$$a\cdot\mathrm{P}\left(X\geq a\right)$$

נעביר אגפים ונקבל את אי השוויון. ■

דוגמה: מחיר כרטיס למשחק קוביה עולה 4 ש`ח. במהלך המשחק מטילים את הקוביה וזוכים בשקלים כמספר שיצא בקוביה.
נשחק 100פעמים ונרצה לחסום את ההסתברות שלא נצא בהפסד. כלומר:

$$\mathrm{P}\left(\sum_{i=1}^{100}X_{i}\geq400\right)$$

כאשר X_i הוא ערך הזכייה במשחק ה־i.
לפי אי־שוויון מרקוב נוכל להסיק:

$$\mathrm{P}\left(\sum_{i=1}^{100}X_{i}\geq400\right)\leq\frac{\mathrm{E}\left(\sum_{i=1}^{100}X_{i}\right)}{400}=\frac{\sum_{i=1}^{100}\mathrm{E}\left(X_{i}\right)}{400}=\frac{100\cdot3.5}{400}=\frac{7}{8}$$

ובאופן שקול נוכל לקבוע:

$$\mathrm{P}\left(\sum_{i=1}^{100}X_{i}<400\right)\geq1-\frac{7}{8}=\frac{1}{8}$$

הערה: אי־שוויון שקול לאי־שוויון מרקוב הוא:

$$\mathrm{P}\left(X \geq t\mathrm{E}\left(X\right)\right) \leq \frac{1}{t}$$

משמעות אי־שוויון זה היא שההסתברות להתרחק מהתוחלת קטנה יותר ככל שהמרחק גדול יותר, כאשר את המרחק מודדים ביחידות התוחלת.

28 אי־שוויון צ'בישב

משפט: בהינתן מ`מ X כלשהו, לכל $a > 0$ מתקיים:

$$\mathrm{P}\left(\left|X - \mathrm{E}\left(X\right)\right| \geq a\right) \leq \frac{\mathrm{Var}\left(X\right)}{a^2}$$

ובאופן שקול:

$$\mathrm{P}\left(\frac{\left|X - \mathrm{E}\left(X\right)\right|}{\mathrm{SD}\left(X\right)} \geq t\right) \leq \frac{1}{t^2}$$

המשמעות היא שההסתברות שמ`מ יהיה רחוק מהתוחלת שלו קטנה ככל שהמרחק ביניהם גדל.

הוכחה: נגדיר מ`מ חדש $\left(X - \mathrm{E}\left(X\right)\right)^2$, ונשים לב שמתקיים לפי הגדרת השונות:

$$\mathrm{E}\left[\left(X - \mathrm{E}\left(X\right)\right)\right]^2 = \mathrm{Var}\left(X\right)$$

כעת בהינתן a נפעיל את אי־שוויון מרקוב עבור a^2:

$$\mathrm{P}\left(\left[X - \mathrm{E}\left(X\right)\right]^2 \geq a^2\right) \leq \frac{\mathrm{E}\left[X - \mathrm{E}\left(X\right)\right]^2}{a^2} = \frac{\mathrm{Var}\left(X\right)}{a^2}$$

נעביר אגפים ונקבל את אי השוויון. ■

המשך הדוגמה: נשים לב שמתקיים עבור $1 \leq i \leq 100$:

$$\mathrm{E}\left(X_i\right) = \tfrac{7}{2}$$

$$\mathrm{Var}\left(X_i\right) = \tfrac{35}{12}$$

ונסמן ונקבל:

$$X = \sum_{i=1}^{100} X_i \Rightarrow \begin{cases} \mathrm{E}\left(X\right) = 350 \\ \mathrm{Var}\left(X\right) = n\mathrm{Var}\left(X_i\right) = 100 \cdot \tfrac{35}{12} \end{cases}$$

(שונות הסכום היא סכום השונויות כי מדובר בסכום של מ`מ ב`ת).
נשתמש בטריק כדי להגיע לביטוי שמכיל את התוחלת 350:

$$\mathrm{P}\left(\sum_{i=1}^{100} X_i \geq 400\right) = \mathrm{P}\left(\sum_{i=1}^{100} X_i - 350 \geq 400 - 350\right) = \mathrm{P}\left(\sum_{i=1}^{100} X_i - 350 \geq 50\right)$$

וכעת נשתמש באי־שוויון צ'בישב כדי לחסום את ההסתברות לרווח:

$$\mathrm{P}\left(\sum_{i=1}^{100} X_i - 350 \geq 50\right) = \frac{1}{2}\cdot\mathrm{P}\left(\left|\sum_{i=1}^{100} X_i - 350\right| \geq 50\right) \leq \frac{1}{2}\cdot\frac{100\cdot\frac{35}{12}}{50^2} = \frac{7}{120}$$

השוויון השני נובע מהסימטריות של המ"מ $X = \sum_{i=1}^{100} X_i$ סביב התוחלת.
נשים לב שקיבלנו חסם משופר משמעותית ביחס לזה שהתקבל באי־שוויון מרקוב, שכן נעזרנו בערך השונות.

דוגמה: מעוניים לאמוד את p שמוגדר כפרופורציית התמיכה במפלגת הליכוד.
נשאל n נדגמים ב"ת האם הם תומכים בליכוד או לא.
נגדיר אינדיקטור לתמיכה בליכוד, $1 \leq i \leq n$:

$$X_i = \begin{cases} 1 & p \\ 0 & 1-p \end{cases}$$

האומד לפרופורציית התומכים בליכוד הוא המ"מ הממוצע $\overline{X} = \frac{1}{n}\cdot\sum_{i=1}^{n} X_i$. נראה מדוע זהו אומד ל־$p$:

$$\mathrm{E}\left(\frac{1}{n}\cdot\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\cdot np = p$$

וכמו־כן:

$$Var\left(\frac{1}{n}\cdot\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\cdot np(1-p) = \frac{p(1-p)}{n} \leq \frac{1}{4n}$$

מפרסמי הסקר מעוניינים שהטעות הסטטיסטית תהיה חסומה על־ידי 3%. כלומר $a = 0.03$.
נדרוש שהוודאות שלא תתרחש טעות של 3% או יותר, תהיה 0.95 לפחות.
ננסח את הדרישה באופן פורמלי:

$$\mathrm{P}\left(\left|\frac{1}{n}\cdot\sum_{i=1}^{n} X_i - p\right| \leq 0.03\right) \geq 1 - \frac{p(1-p)}{n\cdot 0.03^2} \geq 1 - \frac{1}{4n\cdot 0.03^2} \geq 0.95$$

כעת נצטרך למצוא מהו n מספיק גדול שמקיים את הדרישה הזו:

$$1 - \frac{1}{4n\cdot 0.03^2} \geq 0.95$$
$$\Updownarrow$$
$$\frac{1}{4n\cdot 0.03^2} \leq 0.05$$
$$\Updownarrow$$
$$n \geq \frac{1}{4\cdot 0.03^2\cdot 0.05} = 5555.55$$

ולכן כדי לעמוד בדרישות שהצבנו נצטרך לשאול 5556 אנשים לפחות.

29 החוק החלש של המספרים הגדולים

משפט: תהי $\{X_i\}_{i=1}^{n}$ סדרת משתנים מקריים בלתי־תלויים עם תוחלת $\mathrm{E}(X)$ ושונות $\mathrm{Var}(X)$ סופיות ושוות לכולם.
אזי לכל התפלגות בעלת פונקציית הסתברות P ולכל $\varepsilon > 0$ מתקיים:

$$\boxed{\lim_{n\to\infty} \mathrm{P}\left(\left|\frac{1}{n}\cdot\sum_{i=1}^{n} X_i - \mathrm{E}(X)\right| > \varepsilon\right) = 0}$$

הגבול קיין במובן זה שלכל $\varepsilon > 0$ ולכל $\delta > 0$ קיים N, כך שלכל $n > N$ מתקיים:

$$\mathrm{P}\left(\left|\frac{1}{n}\cdot\sum_{i=1}^{n} X_i - \mathrm{E}(X)\right| > \varepsilon\right) < \delta$$

הערה 1: החוק החלש של המספרים הגדולים מתייס לשונות רק בעקיפין - הוא מניח את היותה סופית.

הערה 2: N הוא פונקציה של ε, δ והוא תלוי בהתפלגות.

הוכחה: נפעיל את אי־שוויון צ'בישב על המ`מ $\frac{1}{n}\cdot\sum_{i=1}^{n} X_i$ שתוחלתו היא $\mathrm{E}(X)$ ושונותו היא $\frac{\mathrm{Var}(X)}{n}$:

$$\mathrm{P}\left(\left|\frac{1}{n}\cdot\sum_{i=1}^{n} X_i - \mathrm{E}(X)\right| > \varepsilon\right) \leq \frac{\mathrm{Var}(X)}{n\varepsilon^2} \xrightarrow[n\to\infty]{} 0$$

■

במציאות בדרך כלל לא יודעים מהי התוחלת ולכן לוקחים את ממוצע המדגם במקומה כאומד. החוק אומר שככל שניקח מדגם יותר גדול, כך נקטין את השגיאה שלנו באמידת תוחלת האוכלוסייה באמצעות ממוצע המדגם.

30 שאלות חזרה

1. רוצים לבדוק את אמינות הרישום של שני מיליון תינוקות שנולדו בניו יורק. מתוך שני מליון נדגמו $1,026,000$ בנים. בהנחה כי ההסתברות לבן ולבת שווה וכי הלידות הן בלתי תלויות. האם התוצאה הזו סבירה? כלומר, מהי ההסתברות שמספר הבנים שנדגמו הוא לפחות $1,026,000$? תנו חסם להסתברות זאת בעזרת אי שיוויון מרקוב וצ'יבצ'ב והשוו ביניהם.

2. לחברת ביטוח יש 10,000 מבוטחים שרכשו ביטוח רכב, כל אחד מהם משלם פרמיה שנתית בגובה $r > 0$. סכום התביעה השנתי של מבוטח מהחברה הוא משתנה מקרי בעל תוחלת של 240 ש`ח וסטיית תקן של 800 ש`ח. נניח כי התביעות של המבוטחים ב`ת. בעזרת החוק החלש של המספרים הגדולים רשמו קירוב לרווח הממוצע של החברה עבור כל מבוטח.

חלק VIII
פתרונות

31 סטטיסטיקה תיאורית

1. לפניכם נתונים 13 אורכים שונים **במילימטרים** של גבעולי פרח הכלנית שמדד מר המיולין: $(19, 9, 32, 26, 14, 29, 12, 37, 31, 26, 19, 4, 18)$ חשבו את הממוצע, החציון וסטיית התקן של אורכי הגבעולים. ציירו הסטוגרמה של הנתונים בטווחים של חמישה מילימטרים.

פתרון:

נחשב את הגדלים המבוקשים על פי ההגדרות:

ממוצע אורכי הגבעולים במילמטרים:

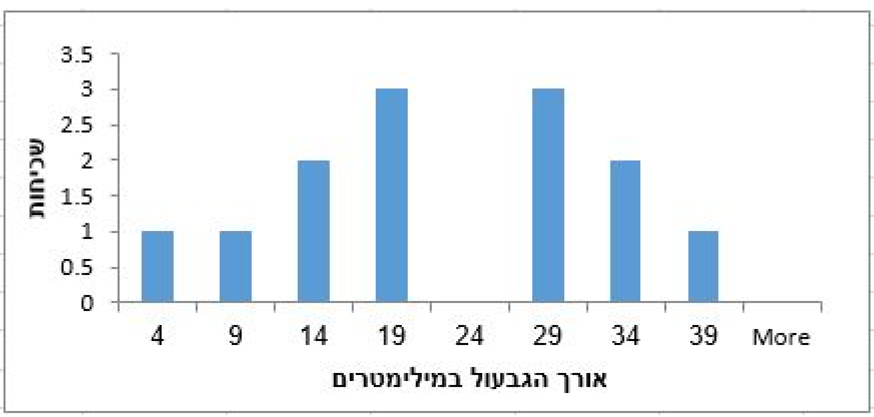
$$\frac{1}{13}(19+9+32+26+14+29+12+37+31+26+19+4+18) = 21.23$$

חציון: נסדר את סדרת המספרים מהקטן לגדול: $(4, 9, 12, 14, 18, 19, 19, 26, 26, 29, 31, 32, 37)$
נזהה שהחציון הינו: 19

סטיית תקן במילימטר מרובע:

$$\sqrt{(19-21.23)^2 + (9-21.23)^2 \ldots\ldots (18-21.23)^2 + (18-21.23)^2} = 9.48$$

הסטוגרמה:



2. בהמשך לנתוני השאלה הקודמת, חשבו את הממוצע, החציון וסטיית התקן של אורכי הגבעולים **בסנטימטרים**, האם ניתן להשתמש בתוצאות השאלה הקודמת?

פתרון:

אין צורך לחשב את הגדלים מחדש, זהו רק שינוי של קנה המידה: 10 מילימטרים = 1 סנטימטר

ממוצע אורכי הגבעולים בסנטימטרים: $\frac{21.23}{10}$

חציון: $\frac{19}{10}$

סטיית תקן בסנטימטר מרובע: $\frac{9.48}{10}$

3. מנהל בית ספר רוצה להוסיף שיעורי תגבור לתלמידי כיתות ו. התקציב של בית הספר מוגבל ומאפשר הוספת שיעורי עזר רק לאחת משלוש הכיתות שבשכבה. במבחן האחרון בחשבון התקבלו התוצאות הבאות: הממוצע בכיתה ו1 הוא 70, החציון 70 וסטיית התקן 5, בכיתה ו2 הממוצע 70החציון 65 וסטיית התקן 5. לאיזה מן הכיתות כדאי לדעתך לתת שיעורי תגבור?

<u>פתרון:</u>

הממוצעים וסטיות התקן בשתי הכיתות זהים ולכן לא ניתן לקבוע על פי מדדים אלו. נסתכל על החציון, החציון בכיתה ו2 נמוך משל ו1 ולכן כדאי את שיעורי התגבור להשקיע בכיתה זו.

4. בנק א' מציע ללקוחותיו תכנית חיסכון לשלוש שנים עם ריבית שנתית של 1% בשנה הראשונה, 5% בשנה השנייה ו־9% בשנה השלישית. בנק ב' מציע ללקוחותיו תכנית חיסכון מקבילה עם ריבית שנתית קבועה של 5% בטענה כי התכנית זהה לזו של בנק א'. האם התוכניות זהות? אם לא, מה צריכה להיות ריבית שנתית קבועה כך שהתוכניות יהיו זהות לחלוטין?

<u>פתרון:</u>

התוכניות אינן זהות. אם נניח שיש לי 100 שקלים אזי בבנק א' לאחר שנה יהיו ברשותי 101 ש"ח, 106.05 לאחר שנתיים ו-115.5945 לאחר שלוש שנים. בבנק ב' לעומת זאת סכומי הכסף שיהיו ברשותי הם בהתאמה 105, 110.25 ו־115.7625. כדי להשיג אותה תשואה בבנק ב' על הריבית השנתית הקבועה להיות: $\sqrt[3]{1.01 \cdot 1.05 \cdot 1.09}$ ז"א ריבית קבועה של 4.49492%

5. הוכח את הטענות הבאות או הבא דוגמא נגדית:

(א) הממוצע תמיד גדול מהנתון המינימאלי

(ב) הממוצע תמיד שווה לחצי מסכום הנתונים המינימאלי והמקסימאלי

(ג) אם הממוצע שווה לחצי מסכום הנתונים המינימאלי והמקסימאלי אז הממוצע והחציון שווים

(ד) אם הממוצע שווה לנתון המקסימאלי אז הנתון המינימאלי והמקסימאלי שווים

<u>פתרון:</u>

(א) לא נכון. כאשר כל הערכים שווים אזי הממוצע שווה לערך המינימלי.

(ב) לא נכון. למשל עבור אוסף המספרים : $1, 1, 1, 1, 1, 3$ – הממוצע הוא 1.333 אולם ממוצע המקסימום והמינימום הוא 2.

(ג) לא נכון. למשל באוסף המספרים:$1, 2, 2, 5, 5$. הממוצע הוא 3 ואילו החציון הוא 2.

(ד) נכון. נסדר את הנתונים בסדר עולה כאשר x_1 הוא הנתון המינימלי ו־ x_n המקסימלי ונניח בשלילה כי $x_1 < x_n$ (המינימום קטן ממש מהמקסימום). אזי

$$\overline{x_n} = \frac{1}{n}\sum_{i=1}^{n} x_i < \frac{1}{n}\sum_{i=1}^{n} x_n = x_n$$

כי $x_i \leq x_n$ לכל i ומההנחה $x_1 < x_n$, קיבלנו כי $\overline{x_n} < x_n$ בניגוד לנתון ולכן $x_1 = x_n$

6. בכיתה ט' 1 ממוצע הציונים בתנ`ך הוא 73, וסטיית התקן היא 18. לימור מכיתה ט' 1 קיבלה 85. בכיתה ט' 2 ממוצע הציונים בתנ`ך הוא 80, וסטיית התקן היא 15. שמעון מכיתה ט' 2 קיבל 98. שמעון: הציון שלי בתנ`ך גבוה משלך, לפיכך אני הצלחתי יותר ממך ! לימור: אז מה, מעמדי בכיתה בתנ`ך גבוה ממעמדך בכיתתך, לכן בעצם, אני הצלחתי יותר ממך! מי צודק? נמקו.

פתרון:

שניהם צודקים לגבי העובדות: שמעון צודק כי מבחינה אבסולוטית, הציון של שלו גבוה מזה של לימור. לימור צודקת כי ציון התקן שלה גבוה מזה של שמעון. ציון התקן של לימוד הוא: $\frac{2}{3} = \frac{85-73}{18}$, ושל שמעון הוא: $\frac{3}{5} = \frac{89-80}{15}$. אם השאלה היא מי יותר קיצוני ביחס להתפלגות הציונים בכיתה שלו (אפילו אם הן לא דומות) אז ציון התקן יכול לענות על שאלה זו.

7. נתונות שתי סדרות שוות אורך של מספרים: $x_1, x_2....x_n$ ו־ $y_1, y_2....y_n$. נגדיר שתי סדרות חדשות באופן הבא: $v_i = x_i y_i$ ו־ $w_i = x_i + y_i$. בכל אחד מהסעיפים הבאים מופיע מדד כלשהו (מרכז או פיזור). קבע אם מדד זה של סדרת ה־w־ים הוא סכום המדדים המתאימים של סדרת ה־x־ים וסדרת ה־y־ים, ואם מדד זה של סדרת ה־v־ים הוא מכפלת המדדים המתאימים של סדרת ה־x־ים וסדרת ה־y־ים. אם תשובתכם היא `כן ־ תמיד`, הוכיחו. אם תשובתכם היא `לא בהכרח תמיד` תנו דוגמה נגדית.

(א) ממוצע

(ב) חציון

(ג) שונות

(ד) סטיית תקן

פתרון:

(א) לגבי w: כן, תמיד. ממוצע הסכומים הוא סכום הממוצעים. הוכחה:

$$\overline{w_n} = \frac{1}{n}\sum_{i=1}^{n} w_i = \frac{1}{n}\sum_{i=1}^{n}(x_i + y_i) = \frac{1}{n}\sum_{i=1}^{n} x_i + \frac{1}{n}\sum_{i=1}^{n} y_i = \overline{x_n} + \overline{y_n}$$

לגבי v: לא בהכרח. נביא דוגמה נגדית:

יהי $n = 2$, והתצפיות הן: $x_1 = 1, x_2 = 3, y_1 = 1, y_2 = 5$ אזי: $v_1 = 1, v_2 = 15$.

הממוצעים הם: $\overline{x_2} = 2, \overline{y_2} = 3, \overline{v_2} = 8$.

בפרט $\overline{v_2} = 8 \neq 2 \cdot 3 = \overline{x_2} \cdot \overline{y_2}$ כלומר: $\overline{v_2} \neq \overline{x_2} \cdot \overline{y_2}$

(ב) גם לגבי w וגם לגבי v: התשובה היא לא בהכרח. נביא דוגמא נגדית:

יהי $n=3$, והתצפיות הן: $x_1=1, x_2=2, x_3=4,\ y_1=8, y_2=4, y_3=6$

אזי: $w_1=9, w_2=6, w_3=10, v_1=8, v_2=8, v_2=24$.

החציונים הם: $med(x)=2, med(y)=6, med(w)=9, med(v)=8$

בפרט: $med(x)\cdot med(y)\neq med(v)$,$med(x)+med(Y)\neq med(w)$

(ג) גם לגבי w וגם לגבי v: התשובה היא לא בהכרח. נביא דוגמא נגדית:

יהי $n=2$, והתצפיות הן: $x_1=0, x_2=2, y_1=0, y_2=4$ אזי: $w_1=0, w_2=6, v_1=0, v_2=8$.

השוניות הן $\mathrm{Var}(x)=1, \mathrm{Var}(y)=4, \mathrm{Var}(w)=9, \mathrm{Var}(v)=16$

בפרט: $\mathrm{Var}(x)\cdot\mathrm{Var}(y)\neq\mathrm{Var}(v)$,$\mathrm{Var}(x)+\mathrm{Var}(y)\neq\mathrm{Var}(w)$

(ד) כיוון שלגבי השונות התשובה אינה נכונה, בפרט היא גם לא תהיה נכונה לגבי סטיית התקן.

8. יהי $x_1, x_2....x_n$ אוסף נתונים כלשהו ויהיו a,b,c,d קבועים. נסמן: $y_i=ax_i+b$ ו־ $w_i=cx_i+d$ עבור $i=1....n$. הראו כי $z_{y_i}=(y_i-\bar{y})/\mathrm{sd}(y)$ ציון התקן של y_i, זהה ל $z_{w_i}=(w_i-\bar{w})/\mathrm{sd}(w)$ ציון התקן של w_i. כלומר, ציוני התקן אינם משתנים עם טרנפורמציות לינאריות.

<u>פתרון:</u>

ראשית נפתח את ציון התקן של w:

$$Z_{w_i}=\frac{(w_i-\overline{w})}{SD(w)}\overset{*}{=}\frac{(cx_i+d-[c\overline{x}+d])}{|c|SD(x)}=\frac{c(x_i-\overline{x})}{|c|SD(x)}=sign(c)Z_{x_i}$$

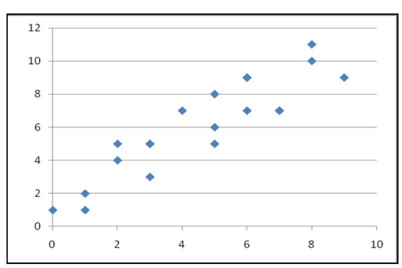
$$sign(c)=\begin{cases}1 & c\geq 0\\ -1 & c<0\end{cases}$$

* לפי תכונות של ממוצע וסטיית תקן

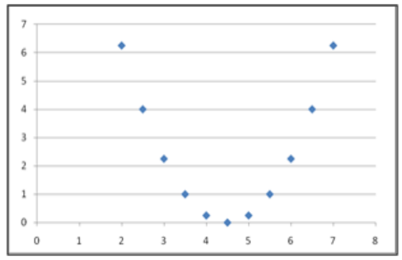
קיבלנו אם כן כי ציון התקן של w זהה לציון התקן של x עד כדי הסימן של c. נוכל לחזור על אותו התהליך גם עבור המשתנה y.

9. נתונות דיאגרמות פיזור. לפני שנותנים את הערכים ומחשבים ערך מספרי למקדם המתאם, מה לדעתכם הוא יצא? מדוע?

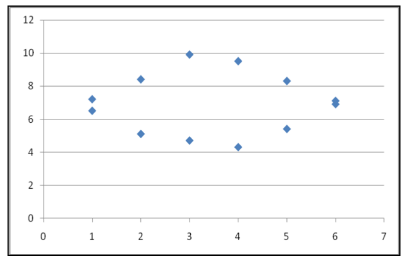
(א)



(ב)



(ג)



פתרון:

(א) אפשר לראות שכאשר X עולה אז גם Y עולה, מכאן שהקשר חיובי. הנקודות נראות סביב קו לינארי, ולכן מקדם המתאם יהיה חיובי ויחסית גבוה

(ב) בדוגמא זאת השונות המשותפת היא קרובה ממש לאפס. מקדם המתאם מעיד על טיב הקשר הלינארי, במקרה שלנו למרות שרואים קשר פונקציונאלי בין X ו־Y (פרבולה) הוא אינו לינארי.

(ג) הפיזור לא מעיד על מגמתיות, ואם היינו מחשבים הפרש מהממוצעים היינו רואים שפעם הסימן חיובי ופעם שלילי ולכן יש קיזוז והשונות המשותפת קרובה לאפס וכך גם מקדם המתאם.

10. להלן זוגות של תצפיות (x_i, y_i) $i = 1...12$:

$$(6,7)(5,8)(5,6)(5,5)(4,7)(3,5)(3,3)(2,5)(2,4)(1,2)(1,1)(0,1)$$

חשבו את השונות המשותפת של x ו־ y.

פתרון:

נחשב את השונות המשותפת על פי הגדרה: $\mathrm{Cov}\,(X,Y) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)$

במקרה שלנו $n = 12$, $\overline{x} = 3.08$ $\overline{y} = 4.5$

$$\mathrm{Cov}\,(x,y) = \frac{1}{12}[(0-3.08)(1-4.5)+(1-3.08)(1-4.5)....+(6-3.08)(7-4.5) = 3.625$$

11. עבור סדרת תצפיות נתון כי $\bar{x} = 3$ $\bar{y} = 3$ ו־ $\bar{xy} = 12$ מהי השונות המשותפת של x ו־ y?

פתרון:

לא ניתן לחשב על פי הגדרה, אך ניתן לראות שמתקיים $\mathrm{Cov}\left(X,Y\right)=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)=$ $\overline{X\cdot Y}-\overline{X}\cdot\overline{Y}$ ולכן $\mathrm{Cov}\left(x,y\right)=12-3\cdot 3=3$

12. יוסי קונה ומוכר בננות מיקרוגלים ומקפיאים. מחיר השוק של כל אחד מאלו נקבע כל יום מחדש, ולכן כל בוקר יוסי צריך ללכת לכל ספק ולשאול אותו מה מחיר המוצר שלו היום. יוסי עצלן ולכן הוא מעדיף ללכת לכמה שפחות ספקים, הוא חושב שבמקום ללכת לכל יבואן (בננות בע`מ, מיקרוגלים בע`מ ומקפיאים בע`מ) ולשאול מה המחיר של המוצר היום, מספיק לו ללכת ליבואן אחד, זאת משום שהוא חושד שהמחיר עולה ויורד בצורה מתואמת ואפשר לנבא את מחיר כל המוצרים בעזרת מחיר של אחד מהם. להלן נתונים של מחירי המוצרים שאסף יוסי במשך עשרה ימי עסקים:

	חברה		
יום עסקים	בננות בע"מ	מיקרוגלים בע"מ	מקפיאים בע"מ
1	31 ש"ח	30 ש"ח	38 ש"ח
2	34 ש"ח	32 ש"ח	36 ש"ח
3	25 ש"ח	40 ש"ח	41 ש"ח
4	23 ש"ח	34 ש"ח	41 ש"ח
5	18 ש"ח	31 ש"ח	32 ש"ח
6	11 ש"ח	28 ש"ח	24 ש"ח
7	4 ש"ח	27 ש"ח	27 ש"ח
8	11 ש"ח	35 ש"ח	20 ש"ח
9	5 ש"ח	40 ש"ח	19 ש"ח
10	7 ש"ח	33 ש"ח	16 ש"ח

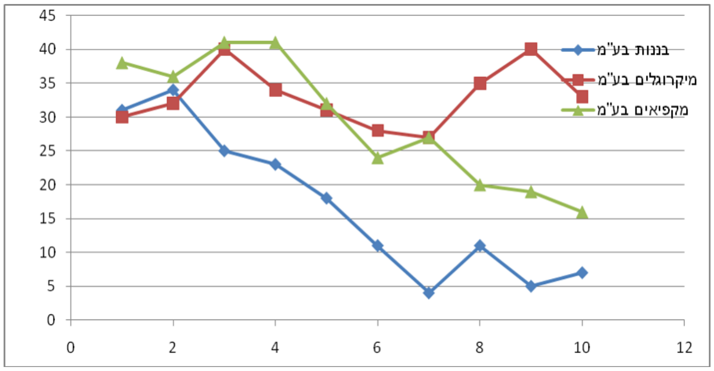
(א) ציירו גרף המתאר את המחיר של כל מוצר כפונקציה של הזמן. האם ניתן להסיק על קשר בין המחירים מתוך הדיאגרמה?

(ב) חשבו את שלושת מקדמי המתאם בין כל זוג מוצרים והחליטו איזה זוג הכי `מבטיח`, נמקו.

(ג) ציירו את דיאגרמות הפיזור של מחירי בננות ומיקרוגלים, בננות ומקפיאים ומיקרוגלים ומקפיאים, האם מקדמי המתאם שחישבתם מתארים נכון את הציורים שקיבלתם?

פתרון:

(א)



מדיאגרמה של מחיר המוצר כפונקציה של הזמן לא ניתן להסיק הרבה. אפשר לראות מגמתיות קלה: כאשר מוצר אחד יורד אז השני גם כן.

(ב) נחשב את שלושת מקדמי המתאם הרלוונטיים: (נסמן: בננות=1, מיקרוגלים=2 מקפיאים=3)

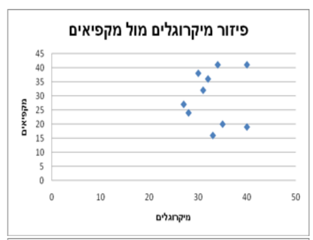
מקדם מתאם בין בננות בע`מ למיקרוגלים בע`מ: $\text{Corr}r_{1,2} = \frac{13}{10.348\cdot 4.219} = 0.029$

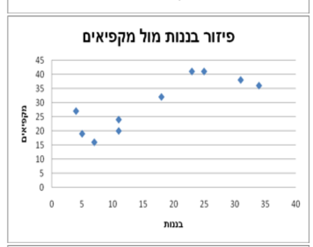
בין בננות בע`מ למקפיאים בע`מ: $\text{Corr}r_{1,2} = \frac{776.4}{10.348\cdot 8.969} = 0.836$

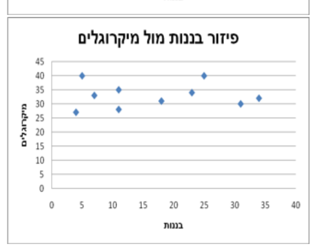
בין מקפיאים בע`מ למיקרוגלים בע`מ: $\text{Corr}r_{1,2} = \frac{5}{8.969\cdot 4.219} = 0.013$

נראה שהזוג הכי `מבטיח` הוא של מחירי הבננות ומחירי המקפיאים, משום שהוא הכי קרוב ל־1 מבין שלושתם

(ג)







ניתן לראות כי בדיאגרמת הפיזור של מקרוגלים ומקפיאים עבור דיאגרמה א' הנקודות מפוזרות באקראיות יחסית ולכן נצפה למתאם קרוב לאפס. בדיאגרמת הפיזור של בננות ומיקרוגלים הפיזור הוא סביב קו שיחסית מקביל לציר ה־X ולכן נצפה גם כאן שהשונות המשותפת תהיה נמוכה. בגרף הפיזור של בננות ומקפיאים ניכר שיש קשר חיובי מכיוון שרואים שכאשר מחיר הבננות עולה המקפיאים המקפיאים עולה איתה, קשר חיובי מצביע כמובן על מתאם חיובי כמו שראינו מהחישוב.

13. סטודנט לתואר ראשון בפסיכולוגיה רצה לבדוק את המתאם בין ציוני הפסכומטרי (טווח ציונים 200־800 ממוצע 500) לציוני מתא`ם (טווח ציונים 50־150 ממוצע 100) הוא החליט שיהיה לו יותר נוח אם ימיר את שתי הסקאלות הציונים לסקאלה אחת. לפיכך, הוא החליט להחסיר 200 נקודות מכל ציון פסיכומטרי, כך שטווח הציונים החדש יהיה 0־600 וממוצע 300, ולהחסיר 50 נקודות מכל ציון מתא`ם ואז להכפילו פי 6 כך שטווח הציונים החדש של המתא`ם יהיה אף הוא 0־600 וממוצע 300. כיצד שינוי הנתונים ישפיע על מקדם המתאם?

<u>פתרון:</u>

מקדם המתאם אינו רגיש לטרנספורמציות ליאריות ולכן לא ישתנה

14. נתונה טבלת ציונים של 21 סטודנטים בשני מקצועות – היסטוריה וכימיה:

סטודנט	כימיה	היסטוריה
1	85	65
2	74	50
3	76	55
4	90	65
5	85	55
6	87	70
7	94	65
8	98	70
9	81	55
10	91	70
11	76	50
12	74	55

(א) ציירו את הנתונים בדיאגרמת פיזור כאשר ציר ה־X הוא הציון בהיסטוריה וציר ה־Y הוא הציון בכימיה, וחשבו את ישר הרגרסיה. ציירו את ישר הרגרסיה של Y על X.

(ב) אהוד קיבל 70 בהיסטוריה, אך החסיר את הבחינה בכימיה בשל מחלה. איזה ציון אופייני לסטודנט כזה לפי הנתונים ?

(ג) א. המורה להיסטוריה החליט לבטל שאלה אחת בבחינה לכן כל תלמיד שיפר את ציונו ב 15% ־ כיצד שינוי זה ישפיע על מקדמי הרגרסיה?

ב. המורה לכימיה גילה כי התלמידים העתיקו לכן הוא הוריד להם 10% מהציון - כיצד שינוי זה ישפיע על מקדמי הרגרסיה?

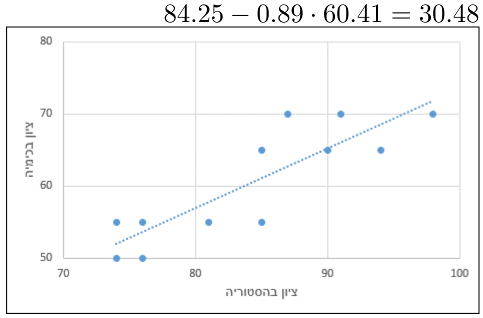
ג. הוכיחו מסקנה כללית לגבי אופן ההשפעה של טרנספורמציה לינארית של הנתונים (הכפלה בקבוע והוספה של קבוע) על מקדמי הרגרסיה. הפרידו למקרים שהטרנספורמציה היא ב־X־ים וב־Y־ים

(ד) חשבו את הקורלציה ואת R בריבוע. בעזרת הקורלציה מצאו את קו הרגרסיה של ציוני התקן של Y על ציוני התקן של X ללא חישובים נוספים.

(ה) חשבו את ציוני התקן של האדם הראשון בקובץ. מה הערך החזוי של אדם זה במונחי ציוני התקן (דהיינו, מה הערך החזוי של ציון התקן בכימיה על פי ציון התקן בהיסטוריה אצל אדם זה). איזו תפועה באה כאן לידי ביטוי?

<u>פתרון:</u>

(א) חישוב מקדמי ישר הרגרסיה: $a = \frac{\text{Cov}(X,Y)}{\text{Var}(X)} = \frac{50.31}{56.07} = 0.89$ $b = \overline{Y} - a\cdot\overline{X} = 84.25 - 0.89\cdot 60.41 = 30.48$



(ב) הציון האופייני לסטודנט שקיבל ציון 70 הוא: $\hat{y} = a\cdot 70 + b = 0.89\cdot 70 + 30.48 = 92.78$

(ג) נענה על השאלה באופן כללי:

מה קורה ל־ a, b כאשר כל x_i מוכפל פי c וערכי y_i לא משתנים?

נסמן בתגים את ערכי ה־x־ים החדשים הממוצע החדש ואת הפרמטרים החדשים.

כלומר $x_i^{'} = cx_i$ $\overline{x_n}' = c\overline{x}_n$

נחשב את המקדמים החדשים:

$$a' = \frac{\sum_{j=1}^{n}(x_i^{'} - \overline{x_n}')(y_i - \overline{y_n})}{\sum_{j=1}^{n}(x_i^{'} - \overline{x_n})^2} = \frac{\sum_{j=1}^{n}(cx_i - c\overline{x_n})(y_i - \overline{y_n})}{\sum_{j=1}^{n}(cx_i - c\overline{x_n})^2} =$$

$$\frac{\sum_{j=1}^{n} c(x_i - \overline{x_n})(y_i - \overline{y_n})}{\sum_{j=1}^{n} c^2(x_i - \overline{x_n})^2} = \frac{1}{c}\cdot\frac{\sum_{j=1}^{n}(x_i - \overline{x_n})(y_i - \overline{y_n})}{\sum_{j=1}^{n}(x_i - \overline{x_n})^2} = \frac{1}{c}\cdot a$$

$$b' = \overline{y_n} - \hat{a}'\overline{x_n}' = \overline{y_n} - \frac{1}{c}\cdot\hat{a}\cdot c\overline{x}_n = \overline{y_n} - \hat{a}\overline{x_n} = b$$

קיבלנו כי כאשר סקלת ה־ X מוכפלת ב־ c השיפוע מחולק ב־ c והחותך אינו משתנה.

מה קורה ל־ a,b כאשר כל y_i מוכפל פי d וערכי x_i לא משתנים?

נסמן בתגים את ערכי ה־y־ים החדשים הממוצע החדש ואת הפרמטרים החדשים.

כלומר $y_i^{'}=cy_i$ $\overline{y_n}^{'}=c\overline{y}_n$

נחשב את המקדמים החדשים:

$$a^{'}=\frac{\sum_{j=1}^{n}\left(x_i-\overline{x_n})(y_i^{'}-\overline{y_n}^{'}\right)}{\sum_{j=1}^{n}\left(x_i^{'}-\overline{x_n})^2\right.}=\frac{\sum_{j=1}^{n}\left(x_i-\overline{x_n})(cy_i-c\overline{y_n}\right)}{\sum_{j=1}^{n}\left(x_i-\overline{x_n})^2\right.}=$$

$$\frac{\sum_{j=1}^{n}c\left(x_i-\overline{x_n})(y_i-\overline{y_n}\right)}{\sum_{j=1}^{n}\left(x_i-\overline{x_n})^2\right.}=\frac{1}{c}\cdot\frac{\sum_{j=1}^{n}\left(x_i-\overline{x_n})(y_i-\overline{y_n}\right)}{\sum_{j=1}^{n}\left(x_i-\overline{x_n})^2\right.}=c\cdot a$$

$$b^{'}=y_i^{'}-\hat{a}^{'}\overline{x_n}^{'}=c\overline{y_n}-c\cdot\hat{a}\cdot c\overline{x}_n=c(\overline{y_n}-\hat{a}\overline{x_n})=c\cdot b$$

קיבלנו כי כאשר סקלת ה־ Y מוכפלת ב־ c השיפוע והחותך מוכפלים ב c.

(ד) לפי הגדרה:

$$\mathrm{Corr}(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\mathrm{SD}(X)\cdot SD(Y)}=\frac{50.31}{7.79\cdot 7.48}=0.86$$

$$R^2=\mathrm{Corr}^2(X,Y)=0.743$$

קו הרגרסיה של ציוני התקן של Y על ציוני התקן של X יהיה: $Z_{\mathrm{y}}=0.86\cdot Z_x$

(ה)

$$Z_{x_1}=\frac{65-60.41}{7.48}=0.61$$

$$\hat{Z}_{y_1}=0.86\cdot 061=0.52$$

ציון התקן החזוי נמוך יותר מציון התקן המסביר, התופעה שבאה כאן לידי ביטוי היא רגרסיה לממוצע

15. רוצים להתאים לנתונים קו רגרסיה מהצורה $y=ax$ (ללא חותך, הקו יוצא מראשית הצירים)

(א) חשבו את פתרון הריבועים הפחותים לבעיה, כלומר, מצאו ערך a המקיים: $\min\sum_{i=1}(y_i-ax_i)^2$

(ב) האם הקו הנ`ל עובר בנקודת הממוצעים?

<u>פתרון:</u>

(א) נגדיר את הפונקציה $g(a)$:

$$g(a) =: \sum_{j=1}^{n} (y_i - ax_i)^2$$

כדי למצוא ערך a המקיים: $min \sum_{i=1} (y_i - ax_i)^2$ נגזור את הפונקציה לפי a ונשווה לאפס:

$$g'(a) = -\sum_{j=1}^{n} 2 \cdot x_i (y_i - ax_i) = 0$$

$$2a \sum_{j=1}^{n} x_i^2 = \sum_{j=1}^{n} 2 \cdot x_i y_i$$

$$\hat{a} = \frac{\sum_{j=1}^{n} x_i y_i}{\sum_{j=1}^{n} x_i^2}$$

(ב) עבור $\hat{a} = \frac{\sum_{j=1}^{n} x_i y_i}{\sum_{j=1}^{n} x_i^2}$ הקו הנ`ל אינו עובר בנקודת הממוצעים. נוכיח על ידי דוגמא פרטית: יהיו $x_1 = 1, x_2 = 2, x_3 = 4,\ y_1 = 8, y_2 = 4, y_3 = 6$
במקרה זה: $\overline{x} = 2.3$ $\overline{y} = 6$ ו־ $a = 1.9$ ומתקיים: $1.9 \cdot 2.3 \neq 6$

16. נזכיר כי עבור הדוגמא של החקלאי וקלחי התירס מ־6.3 קיבלנו את התוצאות הבאות: $\bar{X}_{(25)} = 172.52$ ו־ $SD(X) = 10$. כ"כ רווח הסמך ל μ מתקבל על ידי: $\left[\bar{X}_{(\mathrm{m})} - \frac{\mathrm{SD}(X)}{\sqrt{m}} Z_{1-\frac{\alpha}{2}}, \bar{X}_{(\mathrm{m})} + \frac{\mathrm{SD}(X)}{\sqrt{m}} Z_{1-\frac{\alpha}{2}}\right]$ ולכן:

(א) עבור $\alpha = 0.05$ מתקבל ש: $Z_{1-\frac{\alpha}{2}} = 1.96$ ולכן רווח הסמך ל μ הינו: $\left[172.52 - \frac{10}{\sqrt{25}} 1.96, 172.52 + \frac{10}{\sqrt{25}} 1.96\right] = [168.6, 176.44]$.

(ב) עבור $\alpha = 0.01$ מתקבל ש: $Z_{1-\frac{\alpha}{2}} = 2.58$ ולכן רווח הסמך ל μ הינו: $\left[172.52 - \frac{10}{\sqrt{25}} 2.58, 172.52 + \frac{10}{\sqrt{25}} 2.58\right] = [167.36, 177.68]$.

17. ראשית נחשב מה הוא אורכו של רווח סמך ל μ :

$$\varepsilon = \left(\bar{X}_{(\mathrm{m})} + \frac{\mathrm{SD}(X)}{\sqrt{m}} Z_{1-\frac{\alpha}{2}}\right) - \left(\bar{X}_{(\mathrm{m})} - \frac{\mathrm{SD}(X)}{\sqrt{m}} Z_{1-\frac{\alpha}{2}}\right) = 2 \frac{\mathrm{SD}(X)}{\sqrt{m}} Z_{1-\frac{\alpha}{2}}$$

ומכאן ניתן לבודד את m. אם כך, $m = \left(2\frac{SD(X)}{\varepsilon}Z_{1-\frac{\alpha}{2}}\right)^2$ ועבור הנתונים שלנו נקבל ש: $m = \left(2\frac{1.2}{1.92}1.96\right)^2 = 6.0025 \sim 6$. כלומר גודל המדגם הרצוי הוא לפחות $m = 6$.

18. משרד הבריאות פרסם כי עבור נשים בוגרות מתחת לגיל 51, הקצבה היומית המומלצת של ברזל היא $18mg$. חוקר דגם את רמת הברזל (X) אצל 45 נשים מתחת לגיל 51 במשך יממה. תוצאות המדגם היו: $\bar{X}_{(45)} = 14.68mg$. ונתון: $SD(X) = 4.2mg$. ברמת מובהקות של $\alpha = 0.01$ בדקו את המבחנים הבאים:

(א) $H_0 : \mu = 18mg$ למול $H_1 : \mu \neq 18mg$

(ב) $H_0 : \mu = 18mg$ למול $H_1 : \mu < 18mg$

<u>פתרון:</u>

(א) כזכור אנו נכריע מבחן על סמך סטטיסטי המבחן $P(\mu)$. אם $P(\mu) < \alpha$ נדחה את השערת האפס. עבור מבחן דו צדדי, סטטיסטי המבחן נתון על ידי: $P(\mu) = 2\left(1 - \Phi\left(\frac{|\bar{X}_{(\mathrm{m})} - \mu|}{\mathrm{SD}(X)/\sqrt{\mathrm{m}}}\right)\right)$, ועבור הנתונים שבשאלה, $P(\mu) = 2\left(1 - \Phi\left(\frac{|14.68-18|}{4.2/\sqrt{45}}\right)\right) = 2\left(1 - \Phi(5.3)\right) = 2(1 - 0.99999) = 0 < \alpha = 0.01$, כלומר $P(\mu) = 0 < 0.01 = \alpha$ ולכן החוקר יחליט כנראה לדחות את השערת האפס.

(ב) עבור מבחן חד צדדי, סטטיסטי המבחן נתון על ידי: $P(\mu) = \Phi\left(\frac{\bar{X}_{(\mathrm{m})} - \mu}{\mathrm{SD}(X)/\sqrt{\mathrm{m}}}\right) =$ $\Phi(-5.3) = 0.0000$ ושוב $P(\mu) = 0 < 0.01 = \alpha$ ולכן החוקר יחליט כנראה לדחות את השערת האפס ולקבל את ההשערה האלטרנטיבית.

32 מבוא לתורת הקבוצות ולפונקצית הסתברות

1. מאוכלוסיית הסטודנטים מהחוג לסטטיסטיקה נבחר באקראי סטודנט יחיד נגדיר את המאורעות:

- A ־ הסטודנט קורא ישראל היום.
- B ־ הסטודנט קורא ידיעות אחרונות.
- C ־ הסטודנט קורא הארץ.

(א) השתמשו בפעולות `איחוד`, `חיתוך` ו־`משלים` בלבד, בכדי לבטא בסימונים של קבוצות את המאורעות הבאים.

i. הסטודנט קורא ישראל היום.
ii. הסטודנט קורא ישראל היום וידיעות.
iii. הסטודנט קורא ישראל היום, ידיעות והארץ.
iv. הסטודנט קורא ישראל היום וידיעות, אך אינו קורא הארץ.
v. הסטודנט אינו קורא אף עיתון.
vi. הסטודנט קורא בדיוק שני עיתונים.
vii. הסטודנט קורא לפחות שני עיתונים.

viii. הסטודנט קורא בדיוק עיתון אחד.
ix. הסטודנט קורא לפחות עיתון אחד.
x. הסטודנט אינו `קורא רק הארץ`.

(ב) נסחו במלים את המאורעות הבאים:

i. $B \cup C$
ii. $\overline{A \cup B}$
iii. $(A \cap B \cap \overline{C}) \cup (A \cap \overline{B} \cap C) \cup (A \cap B \cap C)$

(ג) ענו על השאלות הבאות לגבי המאורעות שהוגדרו בסעיף א':

i. האם $i \subseteq ii$? להפך? לא זה ולא זה?
ii. אם $iii \subseteq iv$? להפך? לא זה ולא זה?
iii. האם $viii \subseteq ix$? להפך? לא זה ולא זה?

פתרון:

(א)

i. A
ii. $A \cap B$
iii. $A \cap B \cap C$
iv. $A \cap B \cap \overline{C}$
v. $\overline{A} \cap \overline{B} \cap \overline{C} = \overline{A \cup B \cup C}$
vi. $(A \cap B \cap \overline{C}) \cup (A \cap \overline{B} \cap C) \cup (\overline{A} \cap B \cap C)$
vii. $(A \cap B) \cup (A \cap C) \cup (B \cap C)$
viii. $(A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup (\overline{A} \cap \overline{B} \cap C)$
ix. $A \cup B \cup C$
x. $\overline{(\overline{A} \cap \overline{B} \cap C)} = A \cup B \cup \overline{C}$

(ב)

i. הסטודנט קורא את ידיעות אחרונות, או את הארץ או את שניהם.
ii. הסטודנט לא קורא את ישראל היום או את ידיעות אחרונות.
iii. הסטודנט קורא את כל שלושת העיתונים, או רק את ישראל היום וידיעות אחרונות או רק את ישראל היום והארץ.

(ג)

i. להפך, $ii \subseteq i$, כלומר, $A \cap B \subseteq A$.
ניח שאיבר a נמצא ב-A, אז בפרט הוא נמצא **באחד** מבין השניים, $A \cap B$ או $A \cap \overline{B}$.
לעומת זאת, נניח שאיבר b נמצא ב-$A \cap B$, בפרט הוא נמצא בהכרח ב-A.
הראינו שכל איבר שנמצא ב-$A \cap B$, נמצא ב-A אך לא כל איבר שנמצא ב-A, נמצא ב-$A \cap B$.
ולכן מתקיים $A \cap B \subseteq A$.
ii. לא זה ולא זה. מדובר במאורעות זרים, באחד מהם מופיע חיתוך עם C ובשני חיתוך עם $\overline{C}$, כאשר C ו-$\overline{C}$ זרים.

iii. כן. נראה כי $A \cup B \cup C \supseteq \left(A \cap \overline{B} \cap \overline{C}\right) \cup \left(\overline{A} \cap B \cap \overline{C}\right) \cup \left(\overline{A} \cap \overline{B} \cap C\right)$.

נניח שאיבר a שייך למאורע השמאלי, ז`א שהוא נמצא או ב־A בלבד, או ב־B בלבד, או ב־C בלבד (בגלל החיתוכים עם המאורעות המשלימים). בעוד שהמאורע הימני, $A \cup B \cup C$ הוא איחוד של שלושת המאורעות ולכן כל איבר ששייך אליו, שייך לאחד מהמאורעות, או שניהם, או שלושתם. לכן המאורע הימני מכיל את השמאלי.

2. נתון מרחב המדגם הבא: $\Omega = \{e, f, g, h\}$.

(א) רשמו את כל הקבוצות החלקיות של Ω.

(ב) נגדיר את שלושת הקבוצות הבאות: $A = \{e, f, g\},\ B = \{e, f\},\ C = \{g, h\}$.
הציגו את המאורעות הבאים באמצעות המאורעות הפשוטים של מרחב המדגם:

i. $A \cup B$

ii. $A \cap C$

iii. $\overline{A} \cup B$

iv. $\left(A \cup \overline{B}\right) \cap C$

<u>פתרון:</u>

(א) ל־Ω יש $2^4 = 16$ קבוצות חלקיות:

$$\left\{ \begin{array}{c} \emptyset, (e), (f), (g), (h), \\ (e,f), (e,g), (e,h), (f,g), (f,h), (g,h), \\ (e,f,g), (e,f,h), (e,g,h), (f,g,h), \\ (e,f,g,h) \end{array} \right\}$$

(ב) נרשום את הביטויים בעזרת המאורעות הפשוטים:

i. $A \cup B = \{e, f, g\} \cup \{e, f\} = \{e, f, g\} = A$

ii. $A \cap C = \{e, f, g\} \cap \{g, h\} = \{g\}$

iii. $\overline{A} \cup B = \{h\} \cup \{e, f\} = \{e, f, h\}$

iv. $\left(A \cup \overline{B}\right) \cap C = (\{e, f, g\} \cup \{g, h\}) \cap \{g, h\} = \{e, f, g, h\} \cap \{g, h\} = \{g, h\} = C$

3. הוכיחו באינדוקציה את כללי דה מורגן עבור כל אוסף של מאורעות $A_1, A_2, ..., A_n$, כאשר $n \geq 2$ הוא מספר שלם כלשהו:

(א) $\overline{\cup_{i=1}^{n} A_i} = \cap_{i=1}^{n} \overline{A_i}$

(ב) $\overline{\cap_{i=1}^{n} A_i} = \cup_{i=1}^{n} \overline{A_i}$

<u>פתרון:</u>

(א) עבור $n=2$ ראינו כי מתקיים $\overline{A_1 \cup A_2} = \overline{A_1} \cap \overline{A_2}$.
כעת נניח כי הטענה נכונה עבור $n-1$, כלומר: לכל אוסף מאורעות $A_1, \ldots A_{n-1}$ מתקיים: $\overline{\cup_{i=1}^{n-1} A_i} = \cap_{i=1}^{n-1} \overline{A_i}$.
ניתן להפריד את האיחוד של כל המאורעות לשני חלקים: $\cup_{i=1}^{n} A_i = \left(\cup_{i=1}^{n-1} A_i\right) \cup A_n$. נשים לב כי $\cup_{i=1}^{n-1} A_i$ הוא בעצמו מאורע (כל איחוד או חיתוך של מאורעות הוא מאורע) ולכן ניתן להשתמש בכלל דה מורגן עבור שני מאורעות, ולאחר מכן בהנחת האינדוקציה עבור $n-1$:

$$\overline{\cup_{i=1}^{n} A_i} = \overline{\left(\cup_{i=1}^{n-1} A_i\right) \cup A_n} = \overline{\left(\cup_{i=1}^{n-1} A_i\right)} \cap \overline{A_n} = \cap_{i=1}^{n-1} \overline{A_i} \cap \overline{A_n} = \cap_{i=1}^{n} \overline{A_i}$$

(ב) בעבור $n=2$ ראינו כי מתקיים $\overline{A_1 \cap A_2} = \overline{A_1} \cup \overline{A_2}$.
כעת נניח כי הטענה נכונה עבור $n-1$, כלומר: לכל אוסף מאורעות $A_1, \ldots A_{n-1}$ מתקיים: $\overline{\cap_{i=1}^{n-1} A_i} = \cup_{i=1}^{n-1} \overline{A_i}$.
ניתן להפריד את האיחוד של כל המאורעות לשני חלקים: $\cap_{i=1}^{n} A_i = \left(\cap_{i=1}^{n-1} A_i\right) \cap A_n$. נשים לב כי $\cup_{i=1}^{n-1} A_i$ הוא בעצמו מאורע (כל איחוד או חיתוך של מאורעות הוא מאורע) ולכן ניתן להשתמש בכלל דה מורגן עבור שני מאורעות, ולאחר מכן בהנחת האינדוקציה עבור $n-1$:

$$\overline{\cap_{i=1}^{n} A_i} = \overline{\left(\cap_{i=1}^{n-1} A_i\right) \cap A_n} = \overline{\left(\cap_{i=1}^{n-1} A_i\right)} \cup \overline{A_n} = \cup_{i=1}^{n-1} \overline{A_i} \cup \overline{A_n} = \cup_{i=1}^{n} \overline{A_i}$$

4. נתון מרחב המדגם הבא: $\Omega = \{a,b,c,d,e,f,g,h\}$ קבעו עבור כל אחת מקבוצות המאורעות הבאות האם מדובר בחלוקה של מרחב המדגם. נמקו.

(א) $A=\{a,b,c\}$ $B=\{c,d,e\}$ $C=\{e,f,g,h\}$
(ב) $A=\{a,b,c\}$ $B=\{d,e\}$ $C=\{f,g\}$
(ג) $A=\{a,b,c\}$ $B=\{d,e\}$ $C=\{f,g,h\}$
(ד) $A=\{a,c,h\}$ $B=\{d,e\}$ $C=\{b,f,g\}$

<u>פתרון:</u>

(א) לא חלוקה, הקבוצות A ו־ B לא זרות זו לזו
(ב) לא חלוקה, אחוד של כל הקבוצות אינו Ω
(ג) חלוקה, איחוד של כל הקבוצות הינו Ω, כל הקבוצות זרות זו לזו
(ד) חלוקה, איחוד של כל הקבוצות הינו Ω, כל הקבוצות זרות זו לזו

5. נתון מרחב המדגם הבא: $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, עבורו נתונות ההסתברויות הבאות:
$\mathrm{P}(\{\omega_1\}) = 0.1$, $\mathrm{P}(\{\omega_2\}) = 0.1$, $\mathrm{P}(\{\omega_3\}) = 0.2$, $\mathrm{P}(\{\omega_4\}) = 0.25$, $\mathrm{P}(\{\omega_5\}) = 0.3$, $\mathrm{P}(\{\omega_6\}) = 0.05$.
נגדיר את המאורעות הבאים: $E=\{\omega_1,\omega_2\}$, $F=\{\omega_2,\omega_3\omega_4\}$ ו־ $G=\{\omega_2,\omega_3,\omega_5\}$.

(א) חשבו את ההסתברויות של המאורעות הבאים: $\overline{E}$, $\overline{F}$, $E \cup \overline{F}$, $E \cup G$ ו־ $E \cup F \cup G$.

(ב) בדקו האם מתקיים $\mathrm{P}(E \cup F) = \mathrm{P}(E) + \mathrm{P}(F)$, הסבירו.

(ג) הוכיחו כי P כפי שהוגדרה היא פונקצית הסתברות.

<u>פתרון:</u>

(א) נחשב את ההסתברויות המבוקשות:

$$\mathrm{P}\left(\overline{E}\right) = 1 - \mathrm{P}(E) = 1 - \mathrm{P}\left(\{\omega_1, \omega_2\}\right) = 1 - \mathrm{P}\left(\{\omega_1\}\right) - \mathrm{P}\left(\{\omega_2\}\right) = 1 - 0.1 - 0.1 = 0.8$$
$$\mathrm{P}\left(\overline{F}\right) = \mathrm{P}\left(\{\omega_1, \omega_5, \omega_6\}\right) = \mathrm{P}\left(\{\omega_1\}\right) + \mathrm{P}\left(\{\omega_5\}\right) + \mathrm{P}\left(\{\omega_6\}\right) = 0.1 + 0.3 + 0.05 = 0.45$$
$$\mathrm{P}\left(E \cup \overline{F}\right) = \mathrm{P}\left(\{\omega_1, \omega_2 \omega_5, \omega_6\}\right) = 1 - \mathrm{P}\left(\{\omega_3, \omega_4\}\right) = 1 - \mathrm{P}\left(\{\omega_3\}\right) - \mathrm{P}\left(\{\omega_4\}\right) = 1 - 0.2 - 0.25 = 0.55$$
$$\mathrm{P}\left(E \cup G\right) = P\left(\{\omega_1, \omega_2 \omega_3, \omega_5\}\right) = \mathrm{P}\left(\{\omega_1\}\right) + \mathrm{P}\left(\{\omega_2\}\right) + \mathrm{P}\left(\{\omega_3\}\right) + \mathrm{P}\left(\{\omega_5\}\right) = 0.1 + 0.1 + 0.2 + 0.3 = 0.7$$
$$\mathrm{P}\left(E \cup F \cup G\right) = \mathrm{P}\left(\{\omega_1, \omega_2 \omega_3, \omega_4, \omega_5\}\right) = 1 - \mathrm{P}\left(\{\omega_6\}\right) = 1 - 0.05 = 0.95$$

(ב) נבדוק האם מתקיים השיוויון:

$$\mathrm{P}(E \cup F) = \mathrm{P}\left(\{\omega_1, \omega_2 \omega_3, \omega_4\}\right) = 0.1 + 0.1 + 0.2 + 0.25 = 0.65$$
$$\mathrm{P}(E) + \mathrm{P}(F) = 0.2 + 0.55 = 0.75 \neq \mathrm{P}\left(E \cup F\right)$$

קיבלנו שההסתברות של האיחוד אינו שווה להסתברות של הסכום. תוצאה זאת צפויה כי המאורעות E ו־ F אינם זרים: $E \cap F = \omega_2 \neq \emptyset$.

(ג) נבדוק את שלוש האקסיומות של פונקצית הסתברות:

i. לכל מאורע $A \subseteq \Omega$ ברור כי $\mathrm{P}(A) \geq 0$ כי כל ההסתברויות $P\left(\omega_i\right) \geq 0$ עבור $i = 1, 2, \ldots, 6$.

ii. נוודא שסכום ההסתברויות שווה ל־1:

$$\mathrm{P}\left(\Omega\right) = \mathrm{P}\left(\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}\right)$$

$$= \sum_{i=1}^{6} \mathrm{P}\left(\{\omega_i\}\right) = 0.1 + 0.1 + 0.2 + 0.25 + 0.3 + 0.05 = 1$$

iii. כל זוג מאורעות זרים A ו־B מורכבים מ־ω_i שונים, ולכן ההסתברות של האיחוד $\mathrm{P}\left(A \cup B\right)$ יהיה כמובן שווה לסכום ההסתברויות $\mathrm{P}(A) + \mathrm{P}(B)$.

6. היזכרו בתכונות של פונקצית ההסתברות וענו על השאלות הבאות:

(א) נתונים זוג מאורעות **זרים** $\Omega \supseteq C$ ו־$\Omega \supseteq D$ המקיימים: $\mathrm{P}\left(\overline{C} \cap D\right) = \mathrm{P}\left(C \cap \overline{D}\right)$ ו־ $\mathrm{P}(C) = 0.2$.
חשבו את $\mathrm{P}(C \cup D)$

(ב) יהיו A ו־B שני מאורעות במרחב המדגם. ידוע ש־ $\mathrm{P}\left(B\right) = 0.7$ וכי $\mathrm{P}\left(A\right) = 0.4$.
הוכיחו כי מתקיים: $0.1 \leq \mathrm{P}(A \cap B) \leq 0.4$

(ג) נתונים זוג מאורעות $\Omega \supseteq A$ ו־$\Omega \supseteq B$ המקיימים: $\mathrm{P}\left(\overline{A} \cap \overline{B}\right) = 0.1$, $\mathrm{P}\left(\overline{B}\right) = 0.2$ ו־ $\mathrm{P}\left(A \cap B\right) = 0.6$.
חשבו את $\mathrm{P}(A)$

(ד) ראיתם שלכל זוג מאורעות $A_1, A_2 \subseteq \Omega$ מתקיים: $\mathrm{P}(A_1 \cup A_2) = \mathrm{P}(A_1) + \mathrm{P}(A_2) - \mathrm{P}(A_1 \cap A_2)$,
השתמשו בכך בכדי להראות שלכל שלשה של מאורעות $A_1, A_2, A_3 \subseteq \Omega$ מתקיים:

$$\begin{gathered}\mathrm{P}(A_1 \cup A_2 \cup A_3) \\ = \mathrm{P}(A_1) + \mathrm{P}(A_2) + \mathrm{P}(A_3) - \mathrm{P}(A_1 \cap A_2) \\ -\mathrm{P}(A_1 \cap A_3) - \mathrm{P}(A_2 \cap A_3) + \mathrm{P}(A_1 \cap A_2 \cap A_3)\end{gathered}$$

<u>פתרון:</u>

(א) המאורעות C ו־D זרים, על כן, $C = C \cap \overline{D}$ וגם $D = D \cap \overline{C}$(הסבירו מדוע!)
ומתקיים $\mathrm{P}(C \cap \overline{D}) = \mathrm{P}(C)$ וגם $\mathrm{P}(D \cap \overline{C}) = \mathrm{P}(D)$
ולכן $\mathrm{P}(C \cup D) = \mathrm{P}(C) + \mathrm{P}(D) - \mathrm{P}(C \cap D) = 0.2 + 0.2 - 0 = 0.4$

(ב) ראשית, לכל שני מאורעות A ו־Bמתקיים $A \cap B \subseteq A$ או$A \cap B \subseteq B$ (הסבירו מדוע!)
ולכן מתקיים $\mathrm{P}(A \cap B) \leq \mathrm{P}(A)$ במקרה שלנו ־$\mathrm{P}(A \cap B) \leq 0.4$
שנית, לפי נוסחת האיחוד:

$$\mathrm{P}(A \cap B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cup B) = 0.4 + 0.7 - \mathrm{P}(A \cup B) = 1.1 - \mathrm{P}(A \cup B)$$

$\mathrm{P}(A \cap B) \leq 1$ מתוך הגדרת פונקציית ההסתברות, כלומר: $\mathrm{P}(A \cap B) =$ $1.1 - \mathrm{P}(A \cup B) \leq 1.1$, נקבל: $\mathrm{P}(A \cup B) \geq 0.1$

(ג) בעזרת כללי דה מורגן וההסתברות של המאורע המשלים נסיק כי:

$$\begin{aligned}\mathrm{P}(A \cup B) &= 1 - \mathrm{P}\left(\overline{A \cup B}\right) = 1 - \mathrm{P}\left(\overline{A} \cap \overline{B}\right) = 0.9 \\ \mathrm{P}(B) &= 1 - \mathrm{P}\left(\overline{B}\right) = 0.8\end{aligned}$$

נחשב את ההסתברות המבוקשת על פי נוסחת האיחוד:

$$\begin{aligned}\mathrm{P}(A \cup B) &= \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B) \\ 0.9 &= \mathrm{P}(A) + 0.8 - 0.6 \\ \mathrm{P}(A) &= 0.7\end{aligned}$$

(ד) נסמן $B = A_2 \cup A_3$. לפי הטענה עבור שני מאורעות נקבל,

$$\begin{aligned}\mathrm{P}(A_1 \cup A_2 \cup A_3) &= \mathrm{P}(A_1 \cup B) = \mathrm{P}(A_1) + \mathrm{P}(B) - \mathrm{P}(A_1 \cap B) \\ &= \mathrm{P}(A_1) + \mathrm{P}(A_2 \cup A_3) - \mathrm{P}(A_1 \cap (A_2 \cup A_3))\end{aligned}$$

נחשב בנפרד את שני הביטויים האחרונים:

$$\mathrm{P}\left(A_2 \cup A_3\right) = \mathrm{P}\left(A_2\right) + \mathrm{P}\left(A_3\right) - \mathrm{P}\left(A_2 \cap A_3\right)$$

$$\begin{aligned} \mathrm{P}\left(A_1 \cap \left(A_2 \cup A_3\right)\right) &= \mathrm{P}\left(\left(A_1 \cap A_2\right) \cup \left(A_1 \cap A_3\right)\right) \\ &= \mathrm{P}\left(A_1 \cap A_2\right) + \mathrm{P}\left(A_1 \cap A_3\right) - \mathrm{P}\left(A_1 \cap A_2 \cap A_3\right) \end{aligned}$$

לבסוף נציב את מה שקיבלנו עבור שני ביטויים אלו ונקבל את התוצאה המבוקשת:

$$\begin{gathered} \mathrm{P}\left(A_1 \cup A_2 \cup A_3\right) = \\ \mathrm{P}\left(A_1\right) + \left(\mathrm{P}\left(A_2\right) + \mathrm{P}\left(A_3\right) - \mathrm{P}\left(A_2 \cap A_3\right)\right) \\ - \left(\mathrm{P}\left(A_1 \cap A_2\right) + \mathrm{P}\left(A_1 \cap A_3\right) - \mathrm{P}\left(A_1 \cap A_2 \cap A_3\right)\right) \\ = \mathrm{P}\left(A_1\right) + \mathrm{P}\left(A_2\right) + \mathrm{P}\left(A_3\right) - \mathrm{P}\left(A_1 \cap A_2\right) \\ -\mathrm{P}\left(A_1 \cap A_3\right) - \mathrm{P}\left(A_2 \cap A_3\right) + \mathrm{P}\left(A_1 \cap A_2 \cap A_3\right) \end{gathered}$$

7. בהכנת תמיסה רפואית, עלולות להתרחש תקלות משני סוגים: תקלת `ריכוז` (ריכוז לא נכון) שהסתברותה 0.1, ותקלת `זיהום` (ע`י גורמים שלא אמורים להיות בתמיסה) שהסתברותה 0.04. ההסתברות שלפחות תקלה אחת תתרחש היא 0.13

(א) הגדירו את מרחב המדגם של הניסוי

(ב) מהי ההסתברות ששתי תקלות יתרחשו בו זמנית?

(ג) מהי ההסתברות שלכל היותר תקלה אחת תתרחש?

(ד) מהי ההסתברות שאף תקלה לא תתרחש?

<u>פתרון:</u>

(א) מרחב המדגם: $\Omega = \{(0,0),(0,1),(1,0),(1,1)\}$, כאשר 1 מציין שקרתה תקלה, ו־ 0 שהכל תקין. האיבר הראשון של כל מאורע פשוט במרחב המדגם מתייחס לתקלת ריכוז והאיבר השני מתייחס לתקלת זיהום.

(ב) נגדיר תחילה את המאורעות הבאים:

- $A = \{(1,0),(1,1)\}$ ־ קרתה תקלת ריכוז.
- $B = \{(0,1),(1,1)\}$ ־ קרתה תקלת זיהום.

נתונות לנו ההסתברויות: $\mathrm{P}(A) = 0.1, \mathrm{P}(B) = 0.04, \mathrm{P}\left(A \cup B\right) = 0.13$.

המאורע ששתי התקלות יתרחשו בו זמנית הוא $A \cap B$.

נחשב בעזרת התכונה של פונקצית הסתברות: $\mathrm{P}(A \cap B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cup B)$ את ההסתברות המבוקשת:

$$\mathrm{P}(A \cap B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cup B) = 0.1 + 0.04 - 0.13 = 0.01$$

(ג) המאורע שתתרחש לכל היותר תקלה אחת הוא המשלים של המאורע שיתרחשו שתי תקלות (שחישבנו בסעיף קודם) ולכן:

$$\mathrm{P}\left(\overline{A \cap B}\right) = 1 - \mathrm{P}(A \cap B) = 1 - 0.01 = 0.99$$

(ד) המאורע שלא תתרחש אף תקלה הוא $\overline{A}\cap\overline{B}$. נעזר בכללי דה מורגן:

$$\mathrm{P}\left(\overline{A}\cap\overline{B}\right) = \mathrm{P}\left(\overline{A\cup B}\right) = 1-\mathrm{P}\left(A\cup B\right) = 1-0.13 = 0.87$$

33 קומבינטוריקה

1. כמה סיסמאות ניתן ליצור מהאותיות `סמבטיון` כאשר נדרוש סיסמאות עד אורך של שבע אותיות וניתן להשתמש בכל אות יותר מפעם אחת? (לדוגמא הסיסמאות: ס, סס, ממססב, הן חוקיות)

<u>פתרון:</u>

במילה `סמבטיון` ישנן 7 אותיות. אם רוצים לבחור סיסמא באורך $i \leq 7$ כאשר אפשר לחזור על אותיות יותר מפעם אחת, אז יש 7^i אפשרויות (מדגם סדור עם החזרה, $n = 7$ ו־ $r = i$). ולכן מספר הסיסמאות האפשריות שאורכן הוא עד 7 אותיות הוא הסכום: $\sum_{i=1}^{7} 7^i \overset{*}{=} \frac{7(7^7-1)}{7-1} = 960,799$

*סכום של סדרה הנדסית עם מספר סופי של איברים

2. כמה סיסמאות ניתן ליצור מהאותיות `אנדרלמוסיה` כאשר נדרוש סיסמאות עד אורך של שבע אותיות וניתן להשתמש בכל אות פעם אחת בדיוק?

<u>פתרון:</u>

במילה `אנדרלמוסיה` ישנן 10 אותיות. אם חייבים לבחור כל אות פעם אחת אז אורך הסיסמא חייב להיות כמובן 10. לאות הראשונה יש 10 אפשרויות, ולאחר מכן לאות השנייה נותרו 9 אפשרויות, לשלישית 8 וכך הלאה (מדגם סדור ללא החזרה, $n = 10$ ו־ $r = 10$). לכן מספר הסיסמאות האפשריות: $10! = 3,628,800$

3. על המדף מונחים 22 תבלינים שונים. בכמה דרכים ניתן לתבל אורז לבן ב־3 תבלינים שונים?

<u>פתרון:</u>

ישנן $\binom{22}{3}$ דרכים לבחור 5 בנים מתוך 22 (מדגם לא סדור ללא החזרה, $n = 22$ ו־ $r = 3$)

4. מתוך כיתה של 10 תלמידים ו־7 תלמידות, בוחרים ועד המורכב מ־ 5 תלמידים ו־3 תלמידות. כמה אפשרויות קיימות להרכב הועד הנבחר? (אין חשיבות לסדר בתוך הועד) מה ההסתברות שמיכל ונדב (מועמדים לועד) יבחרו?

<u>פתרון:</u>

ישנן $\binom{10}{5}$ דרכים לבחור 5 בנים מתוך עשרה (מדגם לא סדור ללא החזרה, $n = 10$ ו־ $r = 5$), ו־ $\binom{7}{3}$ דרכים לבחור 3 בנות מתוך שבע (מדגם סדור ללא החזרה, $n = 7$ ו־ $r = 3$). לכן מספר הועדים האפשריים הוא: $\binom{10}{5}\binom{7}{3} = 8,820$.

מספר האפשריות שנדב ומיכל יבחרו הוא כמו מספר האפשרויות לבחור 4 בנים מתוך תשעה ו־2 בנות מתוך שש, לכן הוא $\binom{9}{4}\binom{6}{2} = 2520$ לסיכום ההסתברות המבוקשת היא $\frac{2520}{8820}$

5. בכמה אופנים ניתן לסדר קבוצה של 7 ספרי מתמטיקה שונים, ספרי פיזיקה שונים ו־ 3 ספרי כימיה שונים על מדף? מה תהיה התשובה אם הספרים בכל נושא חייבים להיות סמוכים זה לזה?

<u>פתרון:</u>

מספר הספרים הכולל הוא 12, וכולם שונים. ללא הגבלה על הסידורים יש לספר הראשון יש 12 מקומות, לשני 11 וכך הלאה, (מדגם סדור ללא החזרה, $n = 10$ ו־ $r = 10$), מספר הסידורים האפשריים הוא $12! = 479,001,600$

אם הספרים בכל נושא חייבים להיות סמוכים זה לזה אז ישנם $3!$ סידורים אפשריים של הנושאים על המדף (מדגם סדור ללא החזרה, $n = 3$ ו־ $r = 3$). בנוסף צריך להתחשב במספר הסידורים הפנימיים של כל נושא $7!$ במתמטיקה, $2!$ בפיזיקה ו־$3!$ בכימיה. ולכן מספר האפשרויות הוא: $3! \cdot 7! \cdot 2! \cdot 3! = 362,880$

6. בכמה אופנים שונים אפשר לסדר ארבעה ספרים שונים בשבעה תאים אם בכל תא יש מקום לספר אחד? ומה אם אין הגבלה כזו?

<u>פתרון:</u>

אם בכל תא יש מקום לספר אחד אז לספר הראשון יש 7 תאים אפשריים, לשני 6, לשלישי 5 ולרביעי 4. (מדגם סדור ללא החזרה, $n = 7$ ו־ $r = 4$). סה`כ מספר האפשרויות : $\frac{7!}{(7-4)!} = 7 \cdot 6 \cdot 5 \cdot 4 = 840$

אם אין הגבלה על מספר הספרים בתא אז ניתן להציג את הבעיה הצורה הבאה: צרך לסדר את את הספרות 1, 2, 3, 4 (הספרים) בין 6 מחיצות זהות | (הפרדה למדפים). דוגמא לסידור אפשרי: $2||14|||3|$ בסידור זה ספר 2 נמצא במדף הכי שמאלי, ספרים 4 ו־1 (בסדר הזה) נמצאים במדף השלישי משמאל וספר 3 נמצא במדף השישי משמאל. מספר הסידורים האפשריים של 4 ספרות ושש מחיצות הוא $10!$ אך מכיוון שהמחיצות זהות צריך לחלק במספר הסידורים הפנימיים שלהן: $6!$ לכן מספר הסידורים האפשרי: $\frac{10!}{6!} = 5,040$

7. במרוץ מתחרות שמונה אצניות. מישראל, ירדן, מצרים, עירק, איראן, סוריה, לבנון וטורקיה. האצנית המצרית לא מוכנה לרוץ ליד האצנית הלבנונית. כמה אפשרויות יש לסדר את האצניות באופן כזה? מה ההסתברות כי הטורקיה והלבוננית יוגרלו להיות זו לצד זו?

<u>פתרון:</u>

נחשב את מספר האפשרויות לסדר את האצניות נבחין בין שני סוגי מקרים:

האצנית המצרית נמצאת בקצה המקצה. במקרה זה במקום שלידה יכולה לרוץ כל אחת מהאצניות, חוץ מהאצנית הלבנונית, כלומר 6 אפשריות. מהמקום השלישי נשארו שש אצניות שיכולות לרוץ בכל סדר, ולכן הסידור שלהם יכלול !6 אפשרויות. בסך הכל במקרה כזה יש $6 \cdot 6!$ אפשרויות.

האצנית המצרית יכולה להיות בכל אחד מקצות המקצה ולכן נקבל $6 \cdot 6! \cdot 2$

במקרה השני האצנית המצרית אינה בקצה המקצה. נתחיל לסדר את האצניות שליד האצנית המצרית, מצד ימין יכולות לרוץ 6 אצניות (כולן חוץ מהלבנונית) ומצד שמאל 5 אצניות (כולן חוץ מהלבנונית וממי שכבר שובצה) עכשיו נותר לסדר את כל השאר, 5 אצניות שיכולות לרוץ בכל סדר במקומות שנשארו ולכן נקבל $6 \cdot 5 \cdot 5!$ לאצנית המצרית 6 אפשריות להיות שלא בקצה המקצה ולכן נקבל $6 \cdot 6 \cdot 5 \cdot 5!$ ובסך הכל $30,240$.

נחשב את מספר האפשרויות שבהן האצנית הטורקיה והאצנית הלבנונית רצות זו לצד זו, נבחין הפעם בין שלושה מקרים:

האצנית הטורקיה בקצה המקצה, לידה האצנית הלבנונית, במקרה זה במקום שלידה יכולה לרוץ כל אחת מן האצניות חוץ מן האצנית המצרית, כלומר 5 אפשרויות. מהמקום השלישי נשארו 5 אצניות שיכולות לרוץ בכל סדר ולכן הסידור שלהן יכלול $5!$ אפשרויות. בסך הכל במקרה כזה יש $6 \cdot 6!$ אפשרויות.
האצנית הטורקיה יכולה להיות בכל אחד מקצות המקצה ולכן נקבל $5 \cdot 5! \cdot 2$

במקרה השני האצנית הלבנונית נמצאת בקצה המקצה, לידה האצנית הטורקיה, מהמקום השלישי נשארו 6אצניות שיכולות לרוץ בכל סדר, ולכן הסידור שלהן יכלול $6!$אפשרויות,. האצנית הלבנונית יכולה להיות בכל אחד מקצות המקצה ולכן נקבל $6! \cdot 2$

במקרה השלישי האצנית הלבנונית והטורקיה רצות זו לצד זו, שלא באחד מקצות המקצה, ליד האצנית הלבונית יכולות לרוץ 5 אצניות (כולן חוץ מהמצרית) ובשאר המקומות יכולות לרוץ כל 5 האצניות בכל סדר ולכן נקבל $5 \cdot 5!$, האצנית הלבונית והטורקיה יכולות להחליף בינהן מקומות ולכן בסך הכל נקבל $5{\cdot}5!{\cdot}2$. בסך הכל מספר האפשרויות שהאצנית הטורקיה והלבנונית ירוצו זו לצד זו הוא $5{\cdot}5!{\cdot}2{+}6!{\cdot}2{+}5{\cdot}5!{\cdot}2 =$ 3840. ההסתברות לכך היא $\frac{3840}{30,240}$.

8. בבניין משותף יש ארבע כניסות, ובכל כניסה עשר דירות. בועד של שישה אנשים חשוב שיהיה יצוג לכל כניסה, ומכל דירה יכול להשתתף בועד אדם אחד. לכל אדם בועד תפקיד שונה, בכמה צורות ניתן להרכיב את הועד?

<u>פתרון:</u>

כדי שכל הכניסות תהיינה מיוצגות ישנם שני מקרים. הראשון הוא כאשר לכניסה אחת יש 3 נציגים ולכל אחת משלושת הכניסות הנותרות יש נציג אחד. השני הוא כאשר יש שתי כניסות שלכל אחת מהן ישנם שני נציגים ואילו לכל אחת משתי הכניסות הנותרות יש נציג אחד.

אם כן, תחילה נבחר בועד ורק לאחר מכן נחלק את התפקידים. לכל ועד שנבחר, מספר האפשרויות השונות לחלק את התפקידים הוא $720 = 6!$.

כדי לבחור ועד שבו 3 נציגים מכניסה אחת, יש לבחור את הכניסה ממנה נבחר את שלושת הנציגים, לאחר מכן נבחר את שלושת הנציגים מכניסה זו ואז נציג אחד מכל כניסה אחרת. מספר האפשרויות עבור מקרה זה, אם כן, הוא $\binom{4}{1} \cdot \binom{10}{3} \cdot \binom{10}{1} \cdot \binom{10}{1} \cdot \binom{10}{1} = 4 \cdot 120 \cdot 10^3 = 480,000$

באופן דומה, כאשר ישנן שתי כניסות עם שני נציגים כל אחת, יש לבחור תחילה מיהן שתי הכניסות הללו, אז לבחור שני נציגים מכל אחת מהן ולבסוף נציג אחת מכל אחת

משתי הכניסות הנותרות. לכן במקרה זה נקבל $= \binom{10}{1}\cdot\binom{10}{1}\cdot\binom{10}{2}\cdot\binom{10}{2}\cdot\binom{4}{2}$
$6\cdot 45^2\cdot 10^2 = 1,215,000$

מכאן שמספר הועדים שניתן לבחור כאשר כל כניסה מיוצגת וכאשר התפקידים אינם חשובים הוא $480,000+1,215,000 = 1,695,000$

מספר זה יש להכפיל ב־$720 = 6!$ כדי לקבל את מספר הועדים הרלוונטים עבורם התפקידים כן חשובים. לכן התוצאה המבוקשת היא $1,695,000\cdot 720 = 1,220,400,000$

9. עשרה תלמידים עולים על הסעה מתחנה קבועה. במסלול חמישה בתי ספר וכל תלמיד יורד מההסעה בבית הספר שלו. הנהג רושם כמה נוסעים יורדים בכל תחנה, כמה אפשרויות שונות יש?

<u>פתרון:</u>

יהיה x_i מספר התלמידים שיורדים בבית ספר i. מתקיים $\sum_{i=1}^{5} x_i = 10$ ולכן מספר האפשריות (מדגם לא סדור עם החזרה, $n=5$ ו־ $r=10$) יהיה: $= \binom{5+10-1}{10}$

10. ליגה היא אוסף של קבוצות. בשלב הראשון של התחרות מוגרלות זוגות של קבוצות שמתחרות בינהן. אם בליגה יש עשר קבוצות, כמה אפשרויות שונות יש לסידור קבוצות של השלב הראשון?

<u>פתרון:</u>

נחלק את הבעיה לשלבים, בכל שלב קבוצה אחרת 'תבחר' את בת זוגה למשחק. הקבוצה הראשונה יכולה לבחור כל אחת מתשע הקבוצות. הקבוצה השנייה יכולה לבחור כל אחת משבע הקבוצות שנותרו. וכך עד הסוף. נקבל: $9\cdot 7\cdot 5\cdot 3 = 810$

34 ההסתברות מותנה ואי תלות

1. על פי הגדרה של הסתברות מותנה קיבלנו $\mathrm{P}(A\cap B) = \mathrm{P}(A/B)\mathrm{P}(B)$, הוכיחו באינדוקציה טענה כללית יותר, טענה זו נקראת גם `כלל ההכפלה` או `כלל השרשרת`:

לכל אוסף מאורעות $A_1....A_n$ מתקיים:

$$\mathrm{P}(\cap_{i=1}^{n} A_i) = \mathrm{P}(A_1)\cdot \mathrm{P}(A_2|A)\cdot \mathrm{P}(A_3|A_2\cap A_1)....\mathrm{P}(A_n|A\cap A_{n-2.....}\cap A_1)$$

<u>פתרון:</u>

הטענה נכונה עבור $n=2$, על פי הגדרה של הסתברות מותנית

נניח שהטענה נכונה עבור $n-1$, כלומר מתקיים:

$$\mathrm{P}(\cap_{i=1}^{n-1} A_i) = \mathrm{P}(A_1)\cdot \mathrm{P}(A_2|A_1)\cdot \mathrm{P}(A_3|A_2\cap A_1)....\mathrm{P}(A_{n-1}|A_{n-2}\cap A_{n-2.....}\cap A_1)$$

עלינו להוכיח עבור n:

$$\mathrm{P}(\cap_{i=1}^{n} A_i) = \mathrm{P}(\cap_{i=1}^{n-1} A_i\cap A_n) \stackrel{*}{=} \mathrm{P}(A_n|\cap_{i=1}^{n-1} A_i)\cdot \mathrm{P}(\cap_{i=1}^{n-1} A_i) \stackrel{**}{=}$$

$$\mathrm{P}(A_1)\cdot \mathrm{P}(A_2|A_1)\cdot \mathrm{P}(A_3|A_2\cap A_1)....\mathrm{P}(A_{n-1}|A\cap A_{n-2.....}\cap A_1)\cdot \mathrm{P}(A_n|\cap_{i=1}^{n-1} A)_i)$$

*אם נגדיר $B=\cap_{i=1}^{n-1}A_i)$

** שימוש בהנחת האינדוקציה

2. שישה אנשים משחקים רולטה רוסית. מעבירים אקדח עם כדור אחד אמיתי וחמישה סרק, מקומו של הכדור אינו ידוע. כל אחד בתורו לוחץ על ההדק. האם עדיף לשחק ראשון או אחרון? (הוכיחו את תשובתכם באופן פורמלי, העזרו בכלל השרשרת מהשאלה הקודמת)

<u>פתרון:</u>

סדר השחקנים לא משנה, ההסתברות של כל אחד מהשחקנים לירות את הכדור היא $\frac{1}{6}$.

נראה שזו אכן ההסתברות על ידי שימוש בכלל השרשרת. עבור $i=1,2...6$ נגדיר את המאורעות A_i- השחקן ה- i ירה את הכדור. את המאורע A_i אפשר לרשום גם בצורה הבאה: $A_i=A_i(\cap_{j=1}^{i-1}A_j)$, כלומר זהו המאורע `כל השחקנים לפני שחקן $i$ לא ירו את הכדור והוא כן ירה אותו`. לראשון בתור יש הסתברות של $\frac{1}{6}$לירות את הכדור, ואם לא היה כדור בקנה אז ההסתברות בתור הבא עולה ל- $\frac{1}{5}$ ובתור הבא ל- $\frac{1}{4}$ וכך הלאה. לפי כלל השרשרת אפשר לחשב את ההסתברויות של כל השחקנים, למשל:

$$\mathrm{P}(A_1)=\frac{1}{6}$$

$$\mathrm{P}(A_2)=\mathrm{P}(A_1\cap\overline{A_2})=\mathrm{P}(A_2|\overline{A}_1)=\frac{1}{5}\cdot\frac{5}{6}=\frac{1}{6}$$

$$\mathrm{P}(A_3)=\mathrm{P}(A_3\cap\overline{A_2}\cap\overline{A_1})=\mathrm{P}(A_3|\overline{A_2}\cap\overline{A_1})\cdot\mathrm{P}(\overline{A_2}|A_1)\mathrm{P}(\overline{A_1})=\frac{1}{4}\cdot\frac{4}{5}\cdot\frac{5}{6}=\frac{1}{6}$$

3. יהי מרחב הסתברות עם מרחב מדגם Ω ופונקצית הסתברות P.

הוכח שעבור כל זוג מאורעות $A,B\subseteq\Omega$ בעלי הסתברות חיובית מתקיים: $\mathrm{P}\left(A\cap B|A\right)\geq\mathrm{P}\left(A\cap B|A\cup B\right)$

<u>פתרון:</u>

נפתח את האגף השמאלי לפי ההגדרה של הסתברות מותנית:

$$\mathrm{P}(A\cap B|A)=\frac{\mathrm{P}(A\cap B\cap A)}{\mathrm{P}(A)}=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(A)}$$

ואת האגף הימני באופן דומה:

$$\mathrm{P}(A\cap B|A\cup B)=\frac{\mathrm{P}(\{A\cap B\}\cap\{A\cup B\})}{\mathrm{P}(A\cup B)}\underset{*}{=}\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(A\cup B)}$$

$$\mathrm{P}((A\cap B)\cap(A\cup B))=\mathrm{P}(A\cap B)\Leftarrow(A\cap B)\subseteq(A\cup B)*$$

מתקיים $A\subseteq(A\cup B)$ ולכן $\mathrm{P}(A)\leq\mathrm{P}(A\cup B)$, ואז:

$$\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(A)}\geq\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(A\cup B)}$$

כלומר: $\mathrm{P}(A\cap B|A)\geq \mathrm{P}(A\cap B|A\cup B)$

4. יהי מרחב הסתברות עם מרחב מדגם Ω ופונקצית הסתברות P.

הוכח או הפרך (כלומר, תנו דוגמא נגדית): לכל שלשה של מאורעות $A,B,C\subseteq\Omega$ בעלי הסתברות חיובית מתקיים $\mathrm{P}(A\cap B|A)\geq \mathrm{P}(A\cap B|A\cap C)$

<u>פתרון:</u>

הטענה אינה נכונה.

תחילה נפשט את הביטויים:

$$\begin{aligned}
\mathrm{P}(A\cap B|A) &\overset{?}{\geq} \mathrm{P}(A\cap B|A\cap C)\\
\frac{\mathrm{P}(A\cap B\cap A)}{\mathrm{P}(A)} &\overset{?}{\geq} \frac{\mathrm{P}(A\cap B\cap A\cap C)}{\mathrm{P}(A\cap C)}\\
\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(A)} &\overset{?}{\geq} \frac{\mathrm{P}(A\cap B\cap C)}{\mathrm{P}(A\cap C)}
\end{aligned}$$

נוכל למצוא דוגמא נגדית: נציב למשל - $A=\Omega$:

$$\begin{aligned}
\frac{\mathrm{P}(\Omega\cap B)}{\mathrm{P}(\Omega)} &\overset{?}{\geq} \frac{\mathrm{P}(\Omega\cap B\cap C)}{\mathrm{P}(\Omega\cap C)}\\
\frac{\mathrm{P}(B)}{1} &\overset{?}{\geq} \frac{\mathrm{P}(B\cap C)}{\mathrm{P}(C)}\\
\mathrm{P}(B) &\overset{?}{\geq} \mathrm{P}(B/C)
\end{aligned}$$

הטענה הזו אינה בהכרח נכונה. למשל, עבור הטלת קוביה הוגנת נגדיר את המאורעות הבאים B - יצא 2, C- יצא מספר זוגי. מתקיים: $\mathrm{P}(B)=\frac{1}{6}<\mathrm{P}(B|C)=\frac{1}{3}$

כלומר, עבור המאורעות $A=\Omega$, B - יצא 2, C - יצא מספר זוגי, הטענה לא מתקיימת.

5. יהי מרחב הסתברות עם מרחב מדגם Ω ופונקצית הסתברות P.

הוכח או הפרך: לכל שלשה של מאורעות A B ו- C מתקיים: $\mathrm{P}(A|B\cap C)=\frac{\mathrm{P}(A\cap B|C)}{\mathrm{P}(B|C)}$

<u>פתרון:</u>

הטענה נכונה. לפי ההגדרה של הסתברות מותנה מתקיים: $\mathrm{P}\left(A|B\cap C\right)=\frac{\mathrm{P}(A\cap B\cap C)}{\mathrm{P}(B\cap C)}$, $\mathrm{P}\left(A\cap B|C\right)=\frac{\mathrm{P}(A\cap B\cap C)}{\mathrm{P}(C)}$, $\mathrm{P}\left(B|C\right)=\frac{\mathrm{P}(B\cap C)}{\mathrm{P}(C)}$.

נחשב את המונה והמכנה בעזרת כלל השרשרת/ההכפלה:

$$\begin{aligned}
\mathrm{P}\left(A\cap B\cap C\right) &= \mathrm{P}\left(C\cap\{A\cap B\}\right) = \mathrm{P}\left(C\right)\mathrm{P}\left(A\cap B|C\right)\\
\mathrm{P}\left(B\cap C\right) &= \mathrm{P}(C)\mathrm{P}\left(B|C\right)
\end{aligned}$$

כעת נציב את שני הביטויים בחזרה ונקבל את הטענה שרצינו להוכיח:

$$\mathrm{P}\left(A|B\cap C\right) = \frac{\mathrm{P}\left(A\cap B\cap C\right)}{\mathrm{P}\left(B\cap C\right)} = \frac{\mathrm{P}\left(C\right)\mathrm{P}\left(A\cap B|C\right)}{\mathrm{P}(C)\mathrm{P}\left(B|C\right)} = \frac{\mathrm{P}\left(A\cap B|C\right)}{\mathrm{P}\left(B|C\right)}$$

6. יהי מרחב הסתברות עם מרחב מדגם Ω ופונקצית הסתברות P

הוכח או הפרך: לכל זוג מאורעות A ו־ B מתקיים: $\mathrm{P}(A|B) = \mathrm{P}(B|A)$

פתרון:

הטענה אינה נכונה. נמצא דוגמא נגדית:

מטילים קוביה הוגנת פעמיים, ורושמים את שתי התוצאות. נגדיר את המאורעות הבאים: A ־ סכום התוצאות הוא 8, B ־ התוצאה הראשונה היא 4.
נחשב את שתי ההסתברויות:

$$\begin{aligned} \mathrm{P}(A|B) &= \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6} \\ \mathrm{P}(B|A) &= \frac{\mathrm{P}(B \cap A)}{\mathrm{P}(A)} = \frac{\frac{1}{36}}{\frac{5}{36}} = \frac{1}{5} \end{aligned}$$

ולכן בדוגמא זו: $\mathrm{P}(A|B) \neq \mathrm{P}(B|A)$.

7. באי מסוים שלושה שבטים: שבט דוברי אמת המהווה 40% מאוכלוסיית האי, שבט דוברי שקר המהווה 20% מאוכלוסיית האי ושבט שלישי אשר אנשיו אומרים את האמת בהסתברות 0.6. מה ההסתברות כי אם אשאל תושב מקרי שאלה אחת אקבל תשובה שקרית?

פתרון:

נגדיר מאורעות: A־ פגשתי אדם משבט דוברי האמת, B־ פגשתי אדם משבט דוברי השקר, C־ פגשתי אדם מהשבט השלישי D־ קיבלתי תשובה שקרית.

נתונות לנו ההסתברויות הבאות: $\mathrm{P}(A) = 0.4, \mathrm{P}(B) = 0.2, \mathrm{P}(C) = 0.4$

וגם, $\mathrm{P}(D|A) = 0, \mathrm{P}(D|B) = 1, \mathrm{P}(D|C) = 0.4$

לפי נוסחת ההסתברות השלימה:

$$\mathrm{P}(D) = \mathrm{P}(D|A)\mathrm{P}(A) + \mathrm{P}(D|B)\mathrm{P}(B) + \mathrm{P}(D|C)\mathrm{P}(C)$$

$$= 0 \cdot 0.4 + 1 \cdot 0.2 + 0.4 \cdot 0.4 = 0.36$$

8. אבי ובתיה מוצאים שק עם $2n+1$ מטבעות זהב. בתיה מציעה את המשחק הבא על מנת לקבוע מי יקבל את כל המטבעות
בתיה תטיל $n+1$ מטבעות ואבי את ה־n הנותרים. לאחר מכן הם יספרו את מספרות תוצאות ה`עץ` שכל אחד קיבל, ואם בתיה קיבלה יותר `עץ` אז היא זוכה בכל המטבעות, אחרת אבי זוכה בכל המטבעות. אבי טוען כי המשחק אינו הוגן ואילו בתיה טוענת כי הוא כן, מי צודק?

פתרון:

בתיה צודקת. נוכיח כי ההסתברות שבתיה תזכה היא בדיוק חצי:
נניח שהמשחק מתנהל בצורה הבאה: בכל שלב $i = 1, \ldots, n$ בתיה ואבי מטילים ביחד מטבע ובודקים את התוצאה העדכנית (סך מספר תוצאות העץ של כל אחד עד וכלל לשלב ה־i). בשלב ה־$n+1$ בתיה מטילה את המטבע האחרון שלה ושוב בודקים את התוצאות. נגדיר את המאורעות הבאים:

- A_i ־ לאחר השלב ה־i לבתיה יש יותר `עץ`.
- B_i ־ לאחר השלב ה־i לאבי יש יותר `עץ`.
- C_i ־ לאחר השלב ה־i לשניהם אותו מספר של `עץ`.

ההסתברות אותה אנחנו צריכים לחשב היא של המאורע A_{n+1} ־ לבתיה יש יותר עץ לאחר שהוטלו כל המטבעות. כעת נראה כי בכדי לחשב את ההסתברות מספיק לבחון את מה שקורה בשני שלבים האחרונים. לגבי השלב ה־n אנחנו יכולים לציין את התכונות הבאות:

- אם ידוע כי בתיה הובילה לאחר n שלבים, אז היא בהכרח תוביל לאחר הטלה נוספת של מטבע, כלומר $\mathrm{P}\left(A_{n+1}/|A\right)=1$.
- אם ידוע כי היה תיקו לאחר n שלבים אז בתיה צריכה שההטלה האחרונה תהיה עץ בכדי לנצח: $\mathrm{P}\left(A_{n+1}|C_n\right)=\frac{1}{2}$.
- אם ידוע כי לאחר n שלבים אבי הוביל, אז ההטלה האחרונה לא תשנה שום דבר כי בתיה לכל היותר תשווה את התוצאה: $\mathrm{P}\left(A_{n+1}|B\right)=0$.
- ההסתברות שלאחר n שלבים בתיה מובילה שווה להסתברות שאבי מוביל כי הם הטילו את אותו מספר מטבעות: $\mathrm{P}\left(A_n\right)=\mathrm{P}\left(B_n\right)$. מכך גם נובע: $\mathrm{P}\left(C_n\right)=1-2\mathrm{P}\left(A_n\right)$.

בנוסף נשים לב כי A_n,B_n,C_n מהווים חלוקה של מרחב במדגם (אלה כל האפשרויות בשלב ה־n). לכן אפשר להשתמש בהסתברות השלמה

$$\begin{aligned}
& \mathrm{P}\left(A_{n+1}\right)\\
= & \mathrm{P}\left(A_{n+1}|A\right)\mathrm{P}\left(A_n\right)+\mathrm{P}\left(A_{n+1}|B\right)\mathrm{P}\left(B_n\right)+\mathrm{P}\left(A_{n+1}|C\right)\mathrm{P}\left(C_n\right)\\
= & 1\cdot\mathrm{P}\left(A_n\right)+0\cdot\mathrm{P}\left(B_n\right)+\frac{1}{2}\cdot\mathrm{P}\left(C_n\right)=\mathrm{P}\left(A_n\right)+\frac{1}{2}\left(1-2\mathrm{P}\left(A_n\right)\right)\\
= \tfrac{1}{2} &
\end{aligned}$$

9. לרשותנו מטבע הוגן ושני כדים עם כדורים:

- כד א' מכיל 3 כדורים אדומים ו־2 כחולים
- כד ב' מכיל 2 כדורים אדומים ו־8 כחולים

מטילים את המטבע ואם יצא `עץ` מוציאים כדור מכד א', ואם יצא `פלי` מוציאים כדור מכד ב'. ענו על השאלות הבאות

(א) בהנתן שיצא `עץ`, מה ההסתברות שיצא כדור אדום?
(ב) בהנתן שיצא `פלי` מה ההסתברות שיצא כדור אדום?
(ג) מה ההסתברות שיצא כדור אדום?
(ד) מה ההסתברות שיצא כדור כחול?
(ה) בהנתן שיצא כדור כדור אדום, מה ההסתברות שיצא עץ?
(ו) בהנתן שיצא כדור כדור כחול, מה ההסתברות שיצא עץ?

הערה: לפתרון השאלה הגדירו את המאורעות המתאימים ורשמו באופן מלא את התכונות בהן אתם משתמשים

פתרון:

תחילה נגדיר את המאורעות הבאים: A- המטבע נפל על `עץ`, B- נבחר כדור אדום

המטבע הוגן ולכן $\mathrm{P}(A) = \mathrm{P}(\overline{A}) = \frac{1}{2}$

(א) אם יצא `עץ` אז בוחרים מהכד שבו שלושה כדורים אדומים ושניים כחולים: $\mathrm{P}(B|A) = \frac{3}{3+2} = \frac{3}{5}$

(ב) אם יצא `פלי` אז בוחרים מהכד שבו שני כדורים אדומים ושמונה כחולים: $\mathrm{P}(B|\overline{A}) = \frac{2}{2+8} = \frac{1}{5}$

(ג) לפי נוסחת ההסתברות השלימה:

$$\mathrm{P}(B) = \mathrm{P}(B|A)\mathrm{P}(A) + \mathrm{P}(B|\overline{A})\mathrm{P}(\overline{A}) = \frac{3}{5} \cdot \frac{1}{2} + \frac{1}{5} \cdot \frac{1}{2} = \frac{2}{5}$$

(ד) לפי נוסחת ההסתברות השלימה:

$$\mathrm{P}(\overline{B}) = \mathrm{P}(\overline{B}|A)\mathrm{P}(A) + \mathrm{P}(\overline{B}|\overline{A})\mathrm{P}(\overline{A}) = \frac{2}{5} \cdot \frac{1}{2} + \frac{4}{5} \cdot \frac{1}{2} = \frac{3}{5}$$

כמובן שהיינו יכולים לענות על סעיף זה מייד לאחר סעיף ג:

$$\mathrm{P}(B) = 1 - \mathrm{P}(\overline{B}) = 1 - \frac{2}{5} = \frac{3}{5}$$

(ה) לפי נוסחת בייס: $\mathrm{P}(A|B) = \frac{\mathrm{P}(B/A)\mathrm{P}(A)}{\mathrm{P}(B)} = \frac{\frac{3}{5}\cdot\frac{1}{2}}{\frac{2}{5}} = \frac{3}{4}$

(ו) לפי נוסחת בייס: $\mathrm{P}(A|\overline{B}) = \frac{\mathrm{P}(\overline{B}|A)\mathrm{P}(A)}{\mathrm{P}(\overline{B})} = \frac{\frac{2}{5}\cdot\frac{1}{2}}{\frac{3}{5}} = \frac{1}{3}$

10. עבור זוג מאורעות A ו-B ידוע כי $\mathrm{P}(A) = 0.3$ ו-$\mathrm{P}\left(A \cup B\right) = 0.8$.

(א) הוכיחו כי $\mathrm{P}(B) \geq 0.5$.

(ב) חשבו את $\mathrm{P}(B)$ אם ידוע ש A ו-B זרים.

(ג) חשבו את $\mathrm{P}(B)$ אם ידוע ש A ו-B ב`ת.

פתרון:

(א) לפי נוסחאת האיחוד:

$$\mathrm{P}(B) \quad = \quad \mathrm{P}(A \cup B) - \mathrm{P}(A) + \mathrm{P}(A \cap B) = 0.5 + \mathrm{P}(A \cap B) \geq 0.5$$

(ב) אם A ו־B זרים אז לפי ההגדרה $A\cap B=\phi$, ולכן $\mathrm{P}(A\cap B)=0$. לפי החישוב של הסעיף א' נקבל:

$$\mathrm{P}(B) \;=\; 0.5+\mathrm{P}(A\cap B)=0.5$$

(ג) אם A ו־B ב"ת אז $\mathrm{P}(A\cap B)=\mathrm{P}(A)\cdot\mathrm{P}(B)$, ולפי החישוב של הסעיף א':

$$\begin{gathered}\mathrm{P}(B)=0.5+\mathrm{P}(A\cap B)=0.5+\mathrm{P}(A)\cdot\mathrm{P}(B)=0.5+0.3\mathrm{P}(B)\\ \Leftrightarrow \mathrm{P}(B)\cdot(1-0.3)=0.5\\ \Leftrightarrow \mathrm{P}(B)=\tfrac{5}{7}\end{gathered}$$

11. למשפחה יש $n\geq 2$ ילדים. נגדיר שני מאורעות:

- A ־ למשפחה יש ילדים משני המינים.
- B ־ למשפחה יש לכל היותר בת אחת.

נניח כי ההסתברות ללדת בן או בת היא שווה, וכי הלידות הן ב"ת.

איזה אחת מהטענות הבאות היא הנכונה:

(א) המאורעות A ו־B הינם ב"ת לכל n.

(ב) המאורעות A ו־B הינם תלויים לכל n.

(ג) קיים ערך יחיד של n עבורו A ו־B הינם ב"ת.

(ד) אף אחת מהתשובות א־ג אינה נכונה.

הדרכה: חשבו את $\mathrm{P}(A)$, $\mathrm{P}(B)$ ו־$\mathrm{P}\left(A\cap B\right)$ עבור ערכים קטנים של n ובדקו את הטענות. לאחר מכן נסו למצוא נוסחא כללית להסתברויות הנ"ל וענו בעזרתה על השאלה.

פתרון:

תשובה c נכונה. לכל $n\geq 2$ מרחב המדגם מכיל 2^n מאורעות פשוטים (למשל, עבור $n=2$: $\Omega=\{(girl,girl),(girl,boy),(boy,girl),(boy,\mathrm{boy})\}$)

בעלי הסתברות שווה כל אחד (ההסתברות לבן או לבת זהה)

המשלים של המאורע Aהוא 'למשפחה יש ילדים רק ממין אחד' ולכן: $\mathrm{P}(A)=1-\frac{1}{2^n}-\frac{1}{2^n}=1-\frac{1}{2^{n-1}}$

במאורע Bיש $n+1$ מאורעות פשוטים: 'לא נולדה אף בת' ו־ 'נולדה בת אחת' ולכן, $\mathrm{P}(B)=n\cdot\frac{1}{2^n}+\frac{1}{2^n}=\frac{n+1}{2^n}$

המאורע $A\cap B$הוא המאורע 'נולדה בת אחת בדיוק, ולכן $\mathrm{P}(A\cap B)=n\cdot\frac{1}{2^n}=\frac{n}{2^n}$

בשני החישובים האחרונים השתמשנו בכך שישנן n אפשרויות של בדיוק בת אחת: היא יכולה להיות הראשונה או השנייה או השלישית וכך הלאה עד לאפשרות שהיא הלידה האחרונה.

התנאי לאי תלות הוא $\mathrm{P}(A\cap B)=\mathrm{P}(A)\cdot P(B)$ נציב את הביטויים שקיבלנו: $\left(1-\frac{1}{2^{n-1}}\right)\cdot\frac{n+1}{2^n}=\frac{n}{2^n}$ קל לבדוק ששוויון זה מתקיים רק עבור $n=3$

12. עבור כל אחת מהטענות הבאות ציינו האם היא נכונה או לא. במקרה שכן הוכיחו את הטענה, ובמקרה שלא הציגו דוגמא נגדית להפרכתה:

(א) אם A ו־B ב"ת, אז A ו־$\overline{B}$ ב"ת.

(ב) אם A ו־B ב"ת ו־A ו־C ב"ת, אז A ו־$(B \cup C)$ ב"ת.

(ג) אם A ו־B ב"ת ו־A ו־C ב"ת, אז B ו־C ב"ת.

(ד) אם A ו־$(B \cap C)$ ב"ת, אז A ו־$\left(\overline{B} \cup \overline{C}\right)$ ב"ת.

(ה) אם A בלתי תלוי ב־B, ו־B בלתי תלוי ב־C, אז גם A בלתי תלוי ב־C

<u>פתרון:</u>

(א) הטענה נכונה.

הוכחה: צריך להוכיח כי $\mathrm{P}\left(A \cap \overline{B}\right) = \mathrm{P}(A)\mathrm{P}\left(\overline{B}\right)$ אם A ו־B ב"ת כלומר אם: $\mathrm{P}\left(A \cap B\right) = \mathrm{P}(A)\mathrm{P}(B)$:

$$\begin{aligned}\mathrm{P}\left(A \cap \overline{B}\right) &= \mathrm{P}(A) - \mathrm{P}(A \cap B) = \mathrm{P}(A) - \mathrm{P}(A)P(B) \\ &= \mathrm{P}(A)\left(1 - \mathrm{P}(B)\right) = \mathrm{P}(A)\mathrm{P}\left(\overline{B}\right)\end{aligned}$$

(ב) הטענה אינה נכונה.

נראה דוגמא נגדית: מטילים מטבע הוגן פעמיים. נגדיר את המאורעות הבאים, ונרשום את הסתברותן:

- A ־ יצא "עץ" בדיוק פעם אחת. $\mathrm{P}(A) = \frac{1}{2}$.
- B ־ תוצאת ההטלה הראשונה היא "עץ". $\mathrm{P}(B) = \frac{1}{2}$.
- C ־ תוצאת ההטלה השנייה היא "פלי". $\mathrm{P}(C) = \frac{1}{2}$.

המאורע $A \cap B$ הוא המאורע שיצא "עץ" בראשונה ו־"פלי" בשנייה,

והסתברותו היא $\mathrm{P}(A \cap B) = \frac{1}{2} \cdot \frac{1}{2} = \mathrm{P}(A) \cdot \mathrm{P}(B)$. ולכן A ו־B ב"ת.

בדיוק באותה צורה ניתן לוודא כי גם A ו־C ב"ת.

האיחוד של B ו־C הוא המשלים של המאורע "התקבל פלי בראשונה ועץ בשנייה"

ולכן: $\mathrm{P}\left(B \cup C\right) = 1 - \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}$.

החיתוך $A \cap (B \cup C)$ הוא המאורע "התקבל "עץ" בראשונה ו"פלי" בשנייה". ולכן:

$$\mathrm{P}(A|B \cup C) \quad = \quad \frac{\mathrm{P}(A \cap \{B \cup C\})}{\mathrm{P}(B \cup C)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3} \neq P(A)$$

קיבלנו ש־A אינו ב"ת במאורע $B \cup C$ למרות שהוא ב"ת בכל אחד מהמאורעות באיחוד.

(ג) הטענה אינה נכונה.

נראה דוגמא נגדית: מטילים קובייה הוגנת פעמיים. נגדיר את המאורעות הבאים, ונרשום את הסתברותן:

* A ־ סכום התוצאות הוא 7. $\mathrm{P}(A)=\mathrm{P}\left(\{1,6\},\{6,1\},\{2,5\},\{5,2\},\{3,4\},\{4,3\}\right)=\frac{6}{36}=\frac{1}{6}$.
* B ־ תוצאת ההטלה הראשונה היא 4. $\mathrm{P}(B)=\frac{1}{6}$.
* C ־ תוצאת ההטלה הראשונה היא 3. $\mathrm{P}(C)=\frac{1}{6}$.

נראה כי A ב`ת ב־B וגם ב־C:

$$\begin{aligned}\mathrm{P}(A|B) &= \frac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)}=\frac{\mathrm{P}\left(\{4,3\}\right)}{\frac{1}{6}}=\frac{\frac{1}{36}}{\frac{1}{6}}=\frac{1}{6}=\mathrm{P}(A)\\ \mathrm{P}(A|C) &= \frac{\mathrm{P}(A\cap C)}{\mathrm{P}(C)}=\frac{\mathrm{P}\left(\{3,4\}\right)}{\frac{1}{6}}=\frac{\frac{1}{36}}{\frac{1}{6}}=\frac{1}{6}=\mathrm{P}(A)\end{aligned}$$

לעומת זאת ברור כי B ו־C אינם ב`ת כי הם זרים:

$$\mathrm{P}\left(B\cap C\right) = \mathrm{P}\left(\emptyset\right)=0\neq\mathrm{P}(B)\cdot\mathrm{P}(C)$$

(ד) הטענה נכונה.
הוכחה: אם A ו־$(B\cap C)$ ב`ת אז לפי סעיף א' ניתן להסיק:

$$\begin{aligned}\mathrm{P}\left(A|B\cap C\right) &= \mathrm{P}(A)\\ \Rightarrow\mathrm{P}\left(A|\overline{B\cap C}\right) &= \mathrm{P}(A)\end{aligned}$$

ולפי כלל דה מורגן נקבל:

$$\mathrm{P}\left(A|\overline{B}\cup\overline{C}\right) = \mathrm{P}\left(A|B\cap C\right)=\mathrm{P}(A)$$

ולכן A ו־$\left(\overline{B}\cup\overline{C}\right)$ גם כן ב`ת.

(ה) הטענה אינה נכונה.
נראה דוגמא נגדית: מטילים קוביה הוגנת פעמיים נגדיר את המאורעות: A־ יצא 1 בקוביה הראשונה, B־ יצא 1 בקוביה השנייה, C־ יצא 2 בקוביה הראשונה.
נוכל לשים לב ש־A ו־B בלתי תלויים כי מתקיים $\mathrm{P}(A)=\mathrm{P}(A|B)=\frac{1}{6}$,
ו־B ו־C בלתי תלויים כי מתקיים $\mathrm{P}(C)=\mathrm{P}(C|B)=\frac{1}{6}$
אבל A ו־C אינם בלתי תלויים כי מתקיים ־ $\mathrm{P}(A)=\frac{1}{6}\neq\mathrm{P}(A|C)=0$.

35 משתנים מקריים

1. בבית מלון קטן שלושה חדרים.שלושת אורחי המלון הגיעו למלון שתויים, וכל אחד לקח את אחד משלושת המפתחות באופן מקרי. נגדיר מ`מ X = מספר האורחים (מבין השלושה) שישנו באותו לילה בחדר הנכון.

(א) מצאו את פונקציית ההסתברות ופונקציית ההתפלגות המצטברת של X. ראשית חשבו אילו ערכים יכול המ`מ X לקבל.

(ב) חשבו את ההסתברויות הבאות:

i. $\mathrm{P}(X\geq 2)$

ii. $\mathrm{P}(X \le 1)$

iii. $\mathrm{P}(X < 1)$

<u>פתרון:</u>

(א) מספר האורחים שישנים בחדר הנכון הוא אפס, אחד או שלוש (אם שניים לקחו את המפתח הנכון אז גם השלישי בהכרח לקח את הנכון). כלומר $X \in \{0,1,3\}$.

נגדיר את מרחב המדגם באופן הבא:

$$\Omega = \{(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)\}$$

כאשר האיבר (i,j,k) מסמן את המאורע `אורח 1 קיבל את מפתח $i$, אורח 2 קיבל את מפתח $j$ ואורח 3 קיבל את מפתח $k$`. למשל האיבר $(2,1,3)$ הוא המאורע שאורח 3 לקח את המפתח הנכון ואורחים 1 ו־2 לקחו כל אחד את המפתח של השני.זהו מרחב מדגם אחיד כי כל המפתחות נבחרו באופן אקראי.
נעזר בכך בכדי לחשב את פונקצית ההסתברות של X:

$$\begin{aligned}
\mathrm{P}(X=0) &= \frac{|(2,3,1)\cup(3,1,2)|}{6} = \frac{2}{6} = \frac{1}{3}\\
\mathrm{P}(X=1) &= \frac{|(1,3,2)\cup(2,1,3)\cup(3,2,1)|}{6} = \frac{3}{6} = \frac{1}{2}\\
\mathrm{P}(X=3) &= \frac{|(1,2,3)|}{6} = \frac{1}{6}
\end{aligned}$$

פונקצית ההתפלגות המצטברת תהיה:

$$F_X(x) = \begin{cases} \frac{1}{3} & x=0 \\ \frac{5}{6} & x=1 \\ 1 & x=3 \end{cases}$$

(ב) נחשב את ההסתברויות המבוקשות בעזרת סעיף א':

i. $\mathrm{P}(X \ge 2) = \mathrm{P}(X=3) = \frac{1}{6}$ או $1-F_X(1) = \frac{1}{6}$

ii. $\mathrm{P}(X \le 1) = \mathrm{P}(X=0)+\mathrm{P}(X=1) = \frac{1}{3}+\frac{1}{2} = \frac{5}{6}$ או $F_X(1) = \frac{5}{6}$

iii. $\mathrm{P}(X<1) = \mathrm{P}(X=0) = \frac{1}{3}$ או $F_X(0)=\frac{1}{3}$

2. בכד ארבעה כדורים שחורים, שלושה אדומים ושלושה לבנים.

(א) מוציאים בזה אחר זה 4 כדורים ללא החזרה. יהי X מס' הכדורים האדומים שהוצאו. מהי פונקצית ההסתברות של X ?

(ב) מוציאים בזה אחר זה כדורים ללא החזרה, עד שמתקבל כדור אדום לראשונה. יהי Y מס' הכדורים הכולל שהוצאו. מהי פונקצית ההסתברות של Y ?

<u>פתרון:</u>

(א) בכד ישנם 10 כדורים ובוחרים מתוכם 4, כלומר יש בסך הכל $\binom{10}{4}$ אפשרויות. אנחנו מעוניינים במספר האפשרויות שבהן יש בדיוק x אדומים. תחילה יש לבחור x אדומים מתוך 3 ולאחר מכן $4-x$ נוספים מתוך ה־7 כדורים הלא אדומים. בסה`כ קיבלנו $\binom{3}{x}\cdot\binom{7}{4-x}$ אפשרויות. לכן פונקצית ההסתברות של X היא

$$\mathrm{P}(X=x) = \begin{cases}\frac{\binom{3}{x}\cdot\binom{7}{4-x}}{\binom{10}{4}}, & x=0,1,2,3\\ 0, & otherwise\end{cases}$$

שימו לב כי X מתפלג היפר־גאומטרית.

(ב) תחילה נשים לב כי הערכים ש־Y יכול לקבל הם $\{1,2,3,4,5,6,7,8\}$, כי בניסיון השמיני בוודאות כבר הוצאנו את כל הלבנים והשחורים.

עבור $i\in\{1,\ldots,8\}$ נגדיר את המאורעות A_i ־ הכדור האדום הראשון הוצא בניסיון ה־i. נחשב את פונקצית ההסתברות של Y בעזרת כלל השרשרת (היזכרו בשאלה 1):

$$\begin{aligned}
\mathrm{P}(Y=1) &= P(A_1)=\frac{3}{10}\\
\mathrm{P}(Y=2) &= \mathrm{P}\left(A_2\cap\overline{A_1}\right)=\mathrm{P}\left(A_2|\overline{A_1}\right)\mathrm{P}\left(\overline{A_1}\right)=\frac{3}{9}\cdot\frac{7}{10}=\frac{7}{30}\\
\mathrm{P}(Y=3) &= \mathrm{P}\left(A_3\cap\overline{A_2}\cap\overline{A_1}\right)\\
&= \mathrm{P}\left(A_3|\overline{A_2}\cap\overline{A_1}\right)\mathrm{P}\left(\overline{A_2}|\overline{A_1}\right)\mathrm{P}\left(\overline{A_1}\right)=\frac{3}{8}\cdot\frac{6}{9}\cdot\frac{7}{10}=\frac{7}{40}\\
\mathrm{P}(Y=4) &= \mathrm{P}\big(A_4|\overline{A_3}\cap\overline{A_2}\cap\overline{A_1}\big)\mathrm{P}\left(\overline{A_3}|\overline{A_2}\cap\overline{A_1}\right)\mathrm{P}\left(\overline{A_2}|\overline{A_1}\right)\mathrm{P}\left(\overline{A_1}\right)\\
&= \frac{3}{7}\cdot\frac{5}{8}\cdot\frac{6}{9}\cdot\frac{7}{10}=\frac{1}{8}\\
\mathrm{P}(Y=5) &= \cdots=\frac{3}{6}\cdot\frac{4}{7}\cdot\frac{5}{8}\cdot\frac{6}{9}\cdot\frac{7}{10}=\frac{1}{12}\\
\mathrm{P}(Y=6) &= \cdots=\frac{3}{5}\cdot\frac{3}{6}\cdot\frac{4}{7}\cdot\frac{5}{8}\cdot\frac{6}{9}\cdot\frac{7}{10}=\frac{1}{20}\\
\mathrm{P}(Y=7) &= \cdots=\frac{3}{4}\cdot\frac{2}{5}\cdot\frac{3}{6}\cdot\frac{4}{7}\cdot\frac{5}{8}\cdot\frac{6}{9}\cdot\frac{7}{10}=\frac{1}{40}\\
\mathrm{P}(Y=8) &= \cdots=\frac{3}{3}\cdot\frac{1}{4}\cdot\frac{2}{5}\cdot\frac{3}{6}\cdot\frac{4}{7}\cdot\frac{5}{8}\cdot\frac{6}{9}\cdot\frac{7}{10}=\frac{1}{120}
\end{aligned}$$

3. במבחן אמריקאי 10 שאלות. לכל שאלה 4 תשובות אפשריות. ראובן לא למד כלל, ומנחש את התשובות באופן מקרי. נגדיר מ`מ X = מס' השאלות עליהן ענה ראובן נכון.

(א) כיצד מתפלג X ? כתבו את פונקצית ההסתברות.

(ב) ציון עובר במבחן זה הוא 55. מהי ההסתברות שראובן יעבור את המבחן?

פתרון:

(א) הניחושים הם ב`ת עם הסתברות להצלחה $\frac{1}{4}$ ולכן $X \sim Bin\left(10, \frac{1}{4}\right)$. פונקצית ההסתברות של X היא:

$$\mathrm{P}\left(X=x\right) = \binom{10}{x}\left(\frac{1}{4}\right)^{x}\left(\frac{3}{4}\right)^{10-x}, \; x=0,1,2,\ldots,10$$

(ב) במבחן יש 10 שאלות עם משקל שווה אז ניתן לקבל ציונים רק בכפולות של 10 (אי אפשר לקבל בדיוק 55 למשל). אז בכדי לעבור ראובן חייב לענות על לפחות שש שאלות. ההסתברות לכך היא:

$$\begin{aligned}\mathrm{P}\left(X\geq 6\right) &= \mathrm{P}(X=6)+\mathrm{P}(X=7)+\mathrm{P}(X=8)+\mathrm{P}(X=9)+\mathrm{P}(X=10)\\ &= \binom{10}{6}\left(\frac{1}{4}\right)^{6}\left(\frac{3}{4}\right)^{4}+\binom{10}{7}\left(\frac{1}{4}\right)^{7}\left(\frac{3}{4}\right)^{3}+\binom{10}{8}\left(\frac{1}{4}\right)^{8}\left(\frac{3}{4}\right)^{2}\\ &= +\binom{10}{9}\left(\frac{1}{4}\right)^{9}\left(\frac{3}{4}\right)^{1}+\binom{10}{10}\left(\frac{1}{4}\right)^{10}\left(\frac{3}{4}\right)^{0}\simeq 0.02\end{aligned}$$

4. משך החיים של רכיב מסוג מסוים מתפלג אחיד על פני 12 חודשים (ההסתברות שהוא יתקלקל בכל אחד מהחודשים שווה). במכשיר מותקנים במקביל שלושה מהרכיבים הנ`ל, כאשר המכשיר פועל כל עוד לפחות רכיב אחד תקין. משך החיים של כל רכיב בלתי תלוי במשך החיים של הרכיבים האחרים במכשיר.

(א) מהי ההסתברות שהמכשיר יפעל ללא תקלה יותר מחצי שנה? (רמז: סמנו כמ`מ את מספר הרכיבים התקינים לאחר חצי שנה במכשיר מסוים וחישבו איך הוא מתפלג)

(ב) 10 מכשירים הופעלו באותו רגע (באופן בלתי תלוי אחד בשני), מהי ההסתברות שכעבור חצי שנה יהיו לפחות 9 תקינים?

<u>פתרון:</u>

(א) נתון כי אורך החיים של כל רכיב הוא אחיד על פני שנה שלמה, ולכן ההסתברות שרכיב מסויים יפעל ללא תקלה יותר מחצי שנה, כלומר 7 חודשים או יותר היא: $\frac{1}{2}$.
נסמן ב־X את מספר הרכיבים התקינים לאחר חצי שנה במכשיר מסוים ונשאל את השאלה, מה ההסתבכרות שהמ`מ הזה גדול מ־0, כלומר שלפחות רכיב אחד נשאר פעיל אחרי חצי שנה. במכשיר יש שלושה רכיבים ב`ת עם הסתברות חצי להיות תקינים (תקין ־ 1, לא תקין ־ 0), ומכך ניתן להסיק כי מספר הרכיבים התקינים X הוא מ`מ בינומי: $X \sim Bin\left(3, \frac{1}{2}\right)$ (3 ניסויי ברנולי, כלומר $n=3$ והסתברות של חצי להצלחה). המכשיר פועל אם לפחות רכיב אחד פועל, אז נחשב את ההסתברות לכך לפי פונקצית ההסתברות של ההתפלגות הבינומית:

$$\mathrm{P}\left(X>0\right) = 1-\mathrm{P}(X=0) = 1-\binom{3}{0}\left(\frac{1}{2}\right)^{0}\left(\frac{1}{2}\right)^{3-0} = 1-\frac{1}{8} = \frac{7}{8}$$

(ב) באותו האופן, נסמן ב־Y את מספר המכשירים שפועלים אחרי חצי שנה מתוך 10 שהופעלו יחד. נשים לב שמסעיף קודם ראינו שההסתברות שמכשיר אחד יפעל אחרי חצי שנה היא $\frac{7}{8}$, כעת יש לנו שוב 10 ניסויי ברנולי, כלומר מ`מ שמפולג בינומית: $Y \sim Bin\left(10, \frac{7}{8}\right)$. ההסתברות שלפחות 9 תקינים היא ההסתברות ש־9 תקינים **או** 10 תקינים:

$$\begin{aligned} \mathrm{P}\left(Y \geq 9\right) &= \mathrm{P}\left(Y=9\right)+\mathrm{P}\left(Y=10\right) \\ &= \binom{10}{9}\left(\frac{7}{8}\right)^{9}\left(\frac{1}{8}\right)^{1}+\binom{10}{10}\left(\frac{7}{8}\right)^{10}\left(\frac{1}{8}\right)^{0}=0.64 \end{aligned}$$

5. קבוצה של פיראטים תפסה את הבן של המלך כבן ערובה. הם דרשו תיבה של 50 מטבעות זהב בכדי לשחרר את הבן, אחרת יזרקו אותו לכרישים. המלך שלח לפיראטים תיבה והם צריכים לבדוק אם המטבעות הם באמת מזהב. לפיראטים יש זמן לבדוק רק 5 מטבעות מהתיבה לפני שהם חייבים לברוח. הועלו שתי הצעות לבדיקת התיבה:

(א) פיראט א' הציע לבחור 5 מטבעות בזה אחר זה <u>עם החזרה</u> ולבדוק אם הם מזהב, ואם לפחות אחד הוא מזוייף אז הם יזרקו את בן המלך לכרישים.

(ב) פיראט ב' הציע לבחור 5 מטבעות בזה אחר זה <u>ללא החזרה</u> ולבדוק אם הם מזהב, ואם לפחות אחד הוא מזוייף אז הם יזרקו את בן המלך לכרישים.

המלך שלח לפיראטים תיבה עם 40 מטבעות זהב ו־10 מטבעות מזוייפים. הוכיחו כי ההסתברות שהבן יזרק לכרישים היא גבוהה יותר אם יבצעו את הבדיקה שמציע פיראט ב', לעמות ההצעה של פיראט א'.

<u>פתרון:</u>

נסמן ב־ X את מספר המטבעות המזויפים בבדיקה הראשונה וב־Y את מספר המטבעות המזויפים בבדיקה השנייה.

בבדיקה הראשונה כל מטבע שנבדק מוחזר לתיבה, ולכן ההסתברוץ לבחירת מטבע מזויף שווה ל־$\frac{10}{50}$בכל ניסיון, והנסיונות הם ב`ת. כלומר, מספר המטבעות המזויפים הוא מ`מ בינומי: $X \sim Bin(5, \frac{1}{5})$. ההסתברות לגילוי לפחות מטבע אחד מזויף היא:

$$\mathrm{P}(X \geq 1) = 1-\mathrm{P}(X=0) = 1-(\frac{4}{5})^5 \approx 1-0.328 = 0.672$$

בבדיקה השנייה ההסתברות לגלות מטבע מזויף משתנה בכל ניסיון כי לא מחזירים את המטבעות. ההסתברות לבחור y מטבעות מזויפים מתוך 5 כאשר בתיבה יש 10מטבעות מזויפים בתוך סך הכל של 50 היא לפי ההתפלגות ההיפר גאומטרית:

$$\mathrm{P}(Y=y) = \frac{\begin{pmatrix} 10 \\ y \end{pmatrix}\begin{pmatrix} 40 \\ 5-y \end{pmatrix}}{\begin{pmatrix} 50 \\ 5 \end{pmatrix}}, y \in \{0,1,2,3,4,5\}$$

ההסתברות למציאת מטבע מזויף בשיטה זו היא:

$$P(Y \geq 1) = 1 - P(Y=0) = 1 - \frac{\binom{10}{y}\binom{40}{5-y}}{\binom{50}{5}} \approx 0.689 > 0.672$$

קיבלנו שההסתברות לגילוי מטבע מזויף גדולה יותר בשיטה השנייה.

6. בידי שיכור ופיכח צרור עם n מפתחות זהים למראה, רק אחד מהם פותח את הדלת לביתם. הפיכח מנסה את המפתחות בזה אחר זה בהסתברות שווה **מתוך אלו שלא ניסה עדיין.** לעומת זאת, בכל ניסיון השיכור בוחר מפתח בהסתברות שווה **ללא קשר למה שכבר ניסה.** נגדיר שני משתנים מקריים X- מספר הנסיונות עד שהפיכח יפתח את הדלת, Y- מספר הניסיונות עד שהשיכור יפתח את הדלת. מה ההתפלגות של X ושל Y? עבור $n=5$ מה ההסתברות שכל אחד מהם יצליח לפתוח את הדלת לאחר שני ניסיונות לכל היותר?

פתרון:

התומך של X (הערכים שהמשתנה המקרי Xיכול לקבל בהסתברות גדולה מ־0) הוא $\{1....n\}$. ההסתברות שהמפתח נמצא בכל אחד מהמקומות בסדר אקראי של בדיקה שווה ל- $\frac{1}{n}$, ולכן $X \backsim U(1,n)$. לעומת זאת **בכל** ניסיון של השיכור ההסתברות לבחור במפתח הנכון היא $\frac{1}{n}$ כאשר הנסיונות בלתי תלויים, ולכן $Y \backsim Geo(\frac{1}{n})$. עבור $n=5$ לפי פונקציות ההסתברות המתאימות נחשב:

$$P(X \leq 2) = P(X=1) + P(X=2) = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$$

$$P(Y \leq 2) = P(Y=1) + P(Y=2) = \frac{1}{5}\left(\frac{4}{5}\right)^0 + \frac{1}{5}\left(\frac{4}{5}\right)^1 = \frac{9}{25} < \frac{2}{5}$$

7. מספר הלקוחות המגיעים לבנק בשעה הוא משתנה מקרי פואסון עם פרמטר λ. בבנק יש Kדלפקים ומשך זמן השירות של כל לקוח הוא בדיוק שעה. אם לקוח מגיע ואין דלפק פנוי אז הוא מיד מבלי לקבל שירות. בתחילת היום כל הדלפקים פנויים. חשבו את פונקציית ההסתברות של שני המשתנים המקריים הבאים:

(א) Y- מספר הלקוחות שקיבלו שירות בשעה הראשונה של היום

(ב) Z- מספר הלקוחות שעזבו ללא קבלת שירות בשעה הראשונה של היום

פתרון:

(א) נסמן ב- $X \backsim Pois(\lambda)$את מספר הלקוחות הכולל המגיע לבנק בשעה הראשונה אם הגיעו $X < K$ אז כולם יקבלו שרות, ולכן $P(Y=y) = P(X=y)$ עבור $y < K$. אם $X \geq K$לקוחות יגיעו אז בדיוק Kמתוכם יקבלו שרות, וכל השאר יעזבו ולכן $P(Y=K) = P(X \geq K)$.

לסיכום פונקצית ההסתברות של Y:

$$P(Y=y) = \begin{cases} \frac{e^{-\lambda}\lambda^y}{y!} & 0 \leq y \leq K-1 \\ \sum_{i=K}^{\infty} \frac{e^{-\lambda}\lambda^i}{i!} & y = K \end{cases}$$

(ב) באופן דומה לסעיף הקודם, אם הגיעו $X \leq K$ לקוחות אז כולם יקבלו שרות ואף אחד לא יעזוב, ולכן $\mathrm{P}(Z=0)=\mathrm{P}(X \leq K)$ אם יגיעו $X > K$ אז בדיוק K מתוכם יקבלו שרות ו־ $X-K$ יעזבו, ולכן עבור $z \geq 1$:

$$\mathrm{P}(Z=z)=\mathrm{P}(X-K=z)=\mathrm{P}(X=K+z)$$

לסיכום פונקצית ההסתברות של Z:

$$\mathrm{P}(Z=z))=\begin{cases}\sum_{i=0}^{K}\frac{e^{-\lambda}\lambda^{i}}{i!} & z=0\\ \frac{e^{-\lambda}\lambda^{K+z}}{(K+z)!} & z>0\end{cases}$$

8. בהגרלת הלוטו באי קטן מגרילים מספר זוכה בודד שנבחר באקראי מתוך $\{1,2,...,N\}$ באופן אחיד. בהגרלה משתתפים שני אנשים בלבד, זולו וגולו. כל אחד מהמר על מספר יחיד, והפרס בן S שקלים מחולק באופן שווה בין אלו שניחשו נכונה את המספר הזוכה. זולו וגולו בוחרים את המספר שלהם באקראי (בין 1 ל־N) ובאופן בלתי תלוי.

(א) בכל שבוע יש הגרלה (בשנה יש 52 שבועות). נגדיר X־מספר השבועות בשנה שלפחות אחד מהם (זולו או גולו) זוכה בהגרלה. מהי פונקציית ההסתברות של X?

(ב) מהי תוחלת הרווח של זולו בהגרלה בודדת?

(ג) זולו וגולו רוצים לשפר סיכויים ולכן (לצורך סעיף זה בלבד) הם מחליטים שזולו יבחר במספר 1, גולו יבחר במספר 2 ואם יזכו בפרס כלשהו יחלקו אותו ביניהם. האם השתפרה תוחלת הרווח של כל אחד מהם? הוכיחו.

פתרון:

(א) X הוא מ`מ הסופר את השבועות שבהם לפחות אחד זוכה בהגרלה מכאן שהתומך שלו הוא $\{0,1,...,52\}$. כל אחד מהשבועות הוא ניסוי ברנולי בלתי תלוי בשאר השבועות, שבו `הצלחה` היא שלפחות אחד מהמשתתפים זכה בהגרלה

כלומר, X סופר הצלחות בסדרת ניסויי ברנולי בלתי תלויים ולכן $X \sim Bin(52,p)$.

נחשב את p ־ ההסתברות להצלחה בכל אחד מניסויי הברנולי כלומר, ההסתברות שלפחות אחד מהמשתתפים זוכה בהגרלה:

נסמן ב־ Z זכייה של זולו וב־ G זכייה של גולו, ונשתמש בנוסחת האיחוד.

נשים לב ששני השחקנים בוחרים את המספר שהם מנחשים באופן בלתי תלוי ולכן ההסתברות ששניהם יזכו שווה למכפלת ההסתברויות של כל אחד בנפרד:

כמו כן, נשים לב שהסיכוי של כל אחד מהם לזכות בהגרלה הוא הסיכוי לנחש בדיוק את המספר הזוכה כלומר, $\frac{1}{N}$.

$$\mathrm{P}(Z\cup G)=\mathrm{P}(Z)+\mathrm{P}(G)-\mathrm{P}(G\cap Z)=\mathrm{P}(Z)+\mathrm{P}(G)-\mathrm{P}(G)\cdot\mathrm{P}(Z)$$

$$=\frac{1}{N}+\frac{1}{N}-\frac{1}{N^2}=\frac{2N-1}{N^2}$$

ולכן,

$$X \sim Bin\left(52, \frac{2N-1}{N^2}\right)$$

(ב) כדי לחשב את תוחלת הרווח בהגרלה בודדת, נמצא תחילה את פונקציית ההסתברות:

עבור הגרלה ספציפית כלשהי, נסמן את הרווח של זולו ב־ R. פונקציית ההסתברות של R היא:

r	S	$\frac{S}{2}$	0
$\mathrm{P}(R=r)$	$\frac{N-1}{N^2}$	$\frac{1}{N^2}$	$\frac{N-1}{N}$

כאשר החישוב נעשה כך:

$$\begin{aligned}
\mathrm{P}(R=0) &= \mathrm{P}(zulu\ loses) = \frac{N-1}{N} \\
\mathrm{P}(R=\frac{S}{2}) &= \mathrm{P}(zulu\ wins \cap gulu\ wins) = \frac{1}{N}\cdot\frac{1}{N} = \frac{1}{N^2} \\
\mathrm{P}(R=S) &= \mathrm{P}(zulu\ wins \cap gulu\ loses) = \frac{1}{N}\cdot\frac{N-1}{N} = \frac{N-1}{N^2}
\end{aligned}$$

(בידקו שזו אכן פונקציית הסתברות תקינה)

אם כן, תוחלת הזכייה של זולו בהגרלה בודדת היא: $S\cdot\frac{N-1}{N^2}+\frac{S}{2}\cdot\frac{1}{N}+0\cdot\frac{N-1}{N} = \frac{S(N-0.5)}{N^2}$

(ג) נחשב מחדש את התפלגות הרווח של כל אחד מהם:

r	$\frac{S}{2}$	0
$\mathrm{P}(R=r)$	$\frac{2}{N}$	$\frac{N-2}{N}$

כאשר חישוב ההסתברויות נעשה כך:

$$\begin{aligned}
\mathrm{P}(R=0) &= \mathrm{P}(1\ didn't\ win \cap 2\ didn't\ win) = \frac{N-2}{N} \\
\mathrm{P}(R=\frac{S}{2}) &= \mathrm{P}(1\ won \cup 2\ won) = \mathrm{P}(1\ won) + \mathrm{P}(2\ won) - \mathrm{P}(1\ won \cap 2\ won) = \frac{1}{N}+\frac{1}{N} = \frac{2}{N}
\end{aligned}$$

ולכן תוחלת הרווח החדשה עבור הגרלה בודדת היא: $\frac{S}{2}\cdot\frac{2}{N}+0 = \frac{S}{N}$

נוכל לשים לב ש־ $\frac{S(N-0.5)}{N^2} < \frac{S}{N}$ (מכיוון ש־ $\frac{N-0.5}{N} < 1$), ולכן תוחלת הרווח של כל אחד מהם השתפרה.

9. משה רוצה להשתתף בהגרלה של חבילת נופש, כאשר מספר המשתתפים האחרים בהגרלה אינו קבוע אלא הוא מ`מ שמפולג פואסון $X \sim Poiss(\lambda)$. לכן, הסיכוי לזכות בהגרלה, $Y$, הוא בעצמו משתנה מקרי מפני שהוא פונ' של המ`מ X: $Y = \frac{1}{1+X}$

לשם ההמחשה ניתן לחשוב על כך שההסתברות שמספר המשתתפים חוץ ממשה יהיה 4 היא: $\frac{e^{-\lambda}\lambda^4}{4!}$, לפי פונ' ההסתברות של ההתפלגות הפואסונית. במקרה הזה מתקיים המאורע $\{X=4\}$, כלומר מספר המשתתפים האחרים הוא 4, והסיכוי לזכייה הוא $\frac{1}{1+x}=\frac{1}{1+4}=\frac{1}{5}$ מכיוון שיש סה`כ 5 אנשים שמשתתפים בהגרלה. מהי תוחלת הסיכוי לזכייה בהגרלה?

<u>פתרון:</u>

נשתמש בתכונה של תוחלת של פונקציה של משתנה מקרי $\mathrm{E}(g(x))=\sum_{x=0}^{\infty}\mathrm{P}(X=x)\cdot g(x)$ כאשר $y=g(x)=\frac{1}{1+x}$:

$$\mathrm{E}(Y)=\mathrm{E}\left(\frac{1}{1+X}\right)=\sum_{x=0}^{\infty}\mathrm{P}(X=x)\cdot\frac{1}{1+x}=\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^x}{x!}\cdot\frac{1}{1+x}=$$
$$\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^x}{(x+1)!}=\frac{1}{\lambda}\cdot\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^{x+1}}{(x+1)!}=\frac{1}{\lambda}\cdot\sum_{k=1}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}=$$
$$\frac{1}{\lambda}\cdot\left(\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}-\frac{e^{-\lambda}\lambda^0}{0!}\right)=\frac{1-e^{-\lambda}}{\lambda}$$

10. מושיבים עשר זוגות נשואים סביב שולחן עגול באופן אקראי. מהי תוחלת מספר הנשים שיושבות ליד בעליהן? הדרכה: סמנו את מספר הנשים היושבות ליד בעליהן ב־Y והגדירו את X_i להיות משתנה מקרי שמקבל את הערך 1 אם האישה ה־i יושבת ליד בעלה ו0 אחרת. $i\in\{1,2...10\}$ מה הקשר בין Y לבין $X_1...X_{10}$? השתמשו בתכונת הלינאריות של התוחלת.

<u>פתרון:</u>

נגדירX_i הוא משתנה מקרי ברנולי עם הסתברות p להצלחה ־ `האישה ה־$i$ יושבת ליד בעלה`, כאשר $p=\frac{2}{19}$ (הסבירו למה!)

ומתקיים הקשר: $Y=\sum_{i=1}^{10}X_i$ נחשב עבור כל i: $\mathrm{E}(X_i)=1\cdot\mathrm{P}(X_i=1)+0\cdot\mathrm{P}(X_i=0)=\frac{2}{19}$ ומתכונת הלינאריות של התוחלת נקבל: $\mathrm{E}(Y)=\mathrm{E}(\sum_{i=1}^{10}X_i)=\sum_{i=1}^{10}\mathrm{E}(X_i)=10\cdot\frac{2}{19}=\frac{20}{19}$

11. נתונים שני מטבעות: מטבע 1 עם הסתברות p_1 ל־`עץ` ומטבע 2 עם הסתברות p_2 ל־`עץ`. נגדיר שני משתנים מקריים:

X־ מספר התוצאות `עץ` כאשר בוחרים מטבע באקראי ומטילים אותו פעמיים

Y־ מספר התוצאות `עץ` כאשר בוחרים מטבע באקראי ומטילים אותו פעם אחת ולאחר מכן בוחרים שוב מטבע באופן אקראי ובלתי תלוי בבחירה הקודמת ומטילים פעם נוספת.
האם התוחלת של X גדולה קטנה או שווה לתוחלת של Y? מה לגבי השונות?

<u>פתרון:</u>

נחשב תחילה את פונקציית ההסתברות של X לפי נוסחת ההסתברות השלימה:

$$\mathrm{P}(X=0)=\frac{1}{2}\cdot(1-p_1)^2+\frac{1}{2}\cdot(1-p_2)^2=1+p_1^2+p_2^2-p_1-p_2$$
$$\mathrm{P}(X=1)=\frac{1}{2}\cdot 2\cdot p_1\cdot(1-p_2)+\frac{1}{2}\cdot 2\cdot p_2\cdot(1-p_1)=p_1-p_1^2+P_2-P_2^2$$
$$\mathrm{P}(X=2)=\frac{1}{2}\cdot p_1^2+\frac{1}{2}\cdot P_2^2$$

התוחלת של X:

$$\mathrm{E}(X)=0\cdot(1+p_1^2+p_2^2-p_1-p_2)+1\cdot(p_1-p_1^2+P_2-P_2^2)+2\cdot(\frac{1}{2}\cdot p_1^2+\frac{1}{2}\cdot P_2^2)=p_1+p_2$$

המקרה של Yיותר פשוט, נשים לב כי בכל הטלה ההסתברות ל-`עץ` היא $\frac{1}{2}\cdot p_1+\frac{1}{2}\cdot p_2$, וכי ההטלות בלתי תלויות. לכן Y הוא סכום של ניסויי ברנולי עם אותה הסתברות, כלומר $Y \backsim Bin(2,\frac{1}{2}\cdot p_1+\frac{1}{2}\cdot p_2)$ התוחלת של מ`מ בינומי היא באופן כללי $n\cdot p$ ולכן:

$$\mathrm{E}(Y)=2\cdot(\frac{1}{2}\cdot p_1+\frac{1}{2}\cdot p_2)=p_1+p_2=\mathrm{E}(X)$$

קיבלנו שהתוחלות שוות.

השונות של X:

כדי לחשב את השונות של Xנשתמש בנוסת העבודה לפיה $Var(X)=\mathrm{E}(X^2)-\mathrm{E}^2(X)$ונחשב את המומנט השני של X

$$\mathrm{E}(X^2)=0+1^2\cdot(p_1-p_1^2+P_2-P_2^2)+2^2\cdot(\tfrac{1}{2}\cdot p_1^2+\tfrac{1}{2}\cdot P_2^2)=p_1+p_2+p_1^2+p_2^2$$
$$\Longrightarrow \mathrm{Var}(X)=\mathrm{E}(X^2)-\mathrm{E}^2(X)=p_1+p_2+p_1^2+p_2^2-(p_1+p_2)^2=p_1+p_2-2\cdot p_1\cdot p_2$$

בידקו כי רק כאשר $p_1=p_2$ השונויות של X ושל Y שוות, בכל מקרה אחר השונות של X גדולה משל Y

12. נגדיר:

$$\begin{aligned} X &\sim Bin(2,\frac{1}{5}) \\ Y &= X^3 \\ W &= 8+3X^3 \end{aligned}$$

חשבו את השונות של Y ושל W

<u>פתרון:</u>

השונות של Y: ניעזר בנוסחא $\mathrm{Var}(X)=\mathrm{E}(X^2)-\mathrm{E}^2(X)$
נחשב:

$$\begin{gathered}\mathrm{E}(Y)=E\left(X^3\right)=0^3\cdot\mathrm{P}(X=0)+1^3\cdot\mathrm{P}(X=1)+2^3\cdot\mathrm{P}(X=2)\\ =0\cdot\left(\tfrac{4}{5}\right)^2+1\cdot\left[\begin{pmatrix}2\\1\end{pmatrix}\cdot\tfrac{1}{5}\cdot\tfrac{4}{5}\right]+8\cdot\left(\tfrac{1}{5}\right)^2\\ 0+0.32+0.32=0.64\end{gathered}$$

$$\begin{gathered}\mathrm{E}(Y^2)=\mathrm{E}\left(\left(X^3\right)^2\right)=\mathrm{E}\left(X^6\right)=0^6\cdot\mathrm{P}(X=0)+1^6\cdot\mathrm{P}(X=1)+2^6\cdot\mathrm{P}(X=2)\\ =0\cdot\left(\tfrac{4}{5}\right)^2+1\cdot\left[\begin{pmatrix}2\\1\end{pmatrix}\cdot\tfrac{1}{5}\cdot\tfrac{4}{5}\right]+64\cdot\left(\tfrac{1}{5}\right)^2\\ =0+0.32+2.56=2.88\end{gathered}$$

$$\mathrm{Var}(Y)=2.88-0.64^2=2.4704$$

ההשונות של W: ניעזר ב`לינאריות` השונות: $\mathrm{Var}(aX+b)=a^2\cdot\mathrm{Var}(X)$

נחשב:

$$\mathrm{Var}(W)=\mathrm{Var}(8+3Y)=3^2\cdot\mathrm{Var}(Y)=9\cdot 2.4704=22.2336$$

36 התפלגות משותפת

1. בכד ישנם שלושה כדורים ממוספרים $\{0,1,2\}$. מוציאים שני כדורים עם החזרה ורושמים את התוצאות.
נגדיר את שני המשתנים המקריים הבאים: X ־ המספר הנמוך ביותר שנבחר, Y ־ מספר הכדור הגבוה ביותר שנבחר.

(א) רשמו את פונקצית ההסתברות המשותפת של X ו־Y.

(ב) חשבו את ההתפלגות השולית ואת התוחלות של X ו־Y. האם X ו־Y הם שווי התפלגות?

(ג) חשבו את התוחלות של הסכום $X+Y$ ושל המכפלה XY.

פתרון:

(א) עבור $i=0,1,2$ נגדיר את המאורעות: A_i ־ `התקבל הכדור עם מספר $i$ בהוצאה הראשונה` ו־B_i ־ `התקבל הכדור עם מספר $i$ בההוצאה השנייה`. מכיוון שהניסוי כולל החזרה, אז שתי ההוצאות הם ב`ת, כלומר כל $A_i$ ב`ת בכל B_j וההסתברות להוציא את כל אחד מהכדורים בכל אחת מההוצאות היא $\frac{1}{3}$. נחשב:

$$\begin{aligned}
\mathrm{P}\left(X=0,Y=0\right) &= \mathrm{P}\left(A_0\cap B_0\right)=\mathrm{P}\left(A_0\right)\mathrm{P}\left(B_0\right)=\frac{1}{3}\cdot\frac{1}{3}=\frac{1}{9}\\
\mathrm{P}\left(X=0,Y=1\right) &= \mathrm{P}\left(A_0\cap B_1\right)+\mathrm{P}\left(A_1\cap B_0\right)=\frac{1}{3}\cdot\frac{1}{3}+\frac{1}{3}\cdot\frac{1}{3}=\frac{2}{9}\\
\mathrm{P}\left(X=0,Y=2\right) &= \mathrm{P}\left(A_0\cap B_2\right)+\mathrm{P}\left(A_2\cap B_0\right)=\frac{1}{3}\cdot\frac{1}{3}+\frac{1}{3}\cdot\frac{1}{3}=\frac{2}{9}\\
\mathrm{P}\left(X=1,Y=1\right) &= \mathrm{P}\left(A_1\cap B_1\right)=\frac{1}{3}\cdot\frac{1}{3}=\frac{1}{9}\\
\mathrm{P}\left(X=1,Y=2\right) &= \mathrm{P}\left(A_1\cap B_2\right)+\mathrm{P}\left(A_2\cap B_1\right)=\frac{1}{3}\cdot\frac{1}{3}+\frac{1}{3}\cdot\frac{1}{3}=\frac{2}{9}\\
\mathrm{P}\left(X=2,Y=2\right) &= \mathrm{P}\left(A_2\cap B_2\right)=\frac{1}{3}\cdot\frac{1}{3}=\frac{1}{9}
\end{aligned}$$

בנוסף נשים לב כי $P\left(X=x,Y=y\right)=0$ לכל $x>y$ כי המספר הנמוך ביותר אינו יכול להיות גדול מהמספר הגבוה ביותר.
נסכם את התוצאות בטבלה:

x/y	0	1	2	$P(X=x)$
0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{5}{9}$
1	0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$
2	0	0	$\frac{1}{9}$	$\frac{1}{9}$
$P(Y=y)$	$\frac{1}{9}$	$\frac{3}{9}$	$\frac{5}{9}$	1

(ב) למעשה חישבנו את ההתפלגויות השוליות בטבלה, על ידי סיכום של העמודות והשורות. נרשום את התוצאה כפונקצית הסתברות:

$$\mathrm{P}(X=x) = \begin{cases} \frac{5}{9}, & x=0 \\ \frac{3}{9}, & x=1 \\ \frac{1}{9}, & x=2 \\ 0, & otherwise \end{cases}$$

$$\mathrm{P}(Y=y) = \begin{cases} \frac{1}{9}, & x=0 \\ \frac{3}{9}, & x=1 \\ \frac{5}{9}, & x=2 \\ 0, & otherwise \end{cases}$$

X ו־Y אינם שווי התפלגות. כצפוי, ל־X ישנה הסתברות יותר גבוהה לערך נמוך ול־Y הסתברות יותר גבוהה לערך גבוה.
נחשב את התוחלות בעזרת ההתפלגות השולית:

$$\mathrm{E}(X) = 0\cdot\frac{5}{9}+1\cdot\frac{3}{9}+2\cdot\frac{1}{9}=\frac{5}{9}\approx 0.556$$
$$\mathrm{E}(Y) = 0\cdot\frac{1}{9}+1\cdot\frac{3}{9}+2\cdot\frac{5}{9}=\frac{13}{9}\approx 1.444$$

(ג) את התוחלת של הסכום ניתן לחשב ישירות מהסעיף הקודם, לפי תכונת הליניאריות של התוחלת:

$$\mathrm{E}(X+Y) = \mathrm{E}(X)+\mathrm{E}(Y)=\frac{18}{9}=2$$

את תוחלת המכפלה יש לחשב בעזרת פונקצית ההסתברות המשותפת. נשים לב שכל האיברים שבהם מופיע אפס במכפלה או בהסתברות מתאפסים:

$$\mathrm{E}(XY) = 0+1\cdot1\cdot\frac{1}{9}+1\cdot2\cdot\frac{2}{9}+2\cdot2\cdot\frac{1}{9}=\frac{9}{9}=1$$

2. יהיו X,Y,Z משתנים מקריים, ונתונה פונקצית ההסתברות המשותפת:

$$\mathrm{P}(X=1,Y=2,Z=3)=\frac{1}{4}$$

$$\mathrm{P}(X=2,Y=1,Z=1)=\frac{1}{4}$$

$$\mathrm{P}(X=2,Y=2,Z=1)=\frac{1}{4}$$

$$\mathrm{P}(X=2,Y=3,Z=2)=\frac{1}{4}$$

(א) רשמו את התומכים של X,Y,Z.

(ב) חשבו את ההתפלגויות השוליות ואת התוחלות של Z, Y ו־X, האם הם שווי התפלגות?

(ג) חשבו את התוחלות $\mathrm{E}(XYZ)$ ו־ $\mathrm{E}(XY+XZ+YZ)$.

<u>פתרון:</u>

(א) ניתן לראות מתוך פונקציית ההסתברות המשותפת, כי התומך של X הוא $S_X=\{1,2\}$, התומך של Y הוא $S_Y=\{1,2,3\}$. והתומך של Z הוא $S_Z=\{1,2,3\}$.

(ב) נחשב את ההתפלגויות השוליות כפי שעשינו בשאלה 1, כעת נעבור ישר לרישום פונקציית ההסתברות השולית של כל משתנה ללא טבלה, מכיוון שיש לנו 3 מ`מ:

$$\mathrm{P}\left(X=x\right) = \begin{cases}\frac{1}{4}, & x=1\\ \frac{3}{4}, & x=2\\ 0, & otherwise\end{cases}$$

$$\mathrm{P}\left(Y=y\right) = \begin{cases}\frac{1}{4}, & y=1\\ \frac{1}{2}, & y=2\\ \frac{1}{4}, & y=3\\ 0, & otherwise\end{cases}$$

$$\mathrm{P}\left(Z=z\right) = \begin{cases}\frac{1}{2}, & z=1\\ \frac{1}{4}, & z=2\\ \frac{1}{4}, & z=3\\ 0, & otherwise\end{cases}$$

כפי שניתן לראות, X,Y,Z אינם שווי התפלגות.

נחשב את התוחלות בעזרת ההתפלגויות השוליות:

$$\begin{aligned}\mathrm{E}(X) &= 1\cdot\frac{1}{4}+2\cdot\frac{3}{4}=\frac{7}{4}=1.75\\ \mathrm{E}(Y) &= 1\cdot\frac{1}{4}+2\cdot\frac{1}{2}+3\cdot\frac{1}{4}=2\\ \mathrm{E}(Z) &= 1\cdot\frac{1}{2}+2\cdot\frac{1}{4}+3\cdot\frac{1}{4}=\frac{7}{4}=1.75\end{aligned}$$

(ג) נחשב את התוחלת של המכפלה של XYZ
באמצעות חישוב תוחלת של פונקצית משתנים מקריים כך ש־ $g(x,y,z)=xyz$.
ניזכר כי התוחלת של פונקציה של מ`מ מחושבת במקרה זה כך:

$$\mathrm{E}\left(g\left(X,Y,Z\right)\right) = \sum_{x\in S_X}\sum_{y\in S_Y}\sum_{z\in S_Z} g(x,y,z)\cdot \mathrm{P}\left(X=x,Y=y,Z=z\right)$$

כאשר, כאשר S_X,S_Y,S_Z הם התומכים שחושבו בסעיף א'.
נחשב את התוחלת על ידי סכימה של המכפלות האפשריות כפול ההסתברויות המתאימות, נשים לב שישנן רק 4 מכפלות כאלה, לפי פונקציית ההסתברות משותפת ונקבל:

$$\begin{gathered}\mathrm{E}(XYZ)=1\cdot 2\cdot 3\cdot \mathrm{P}(X=1,Y=2,Z=3)+2\cdot 1\cdot 1\cdot \mathrm{P}(X=2,Y=1,Z=1)\\ +2\cdot 2\cdot 1\cdot \mathrm{P}(X=2,Y=2,Z=1)+2\cdot 3\cdot 2\cdot \mathrm{P}(X=2,Y=3,Z=2)\\ =\tfrac{6}{4}+\tfrac{2}{4}+\tfrac{4}{4}+\tfrac{12}{4}=6\end{gathered}$$

באותה צורה ניתן לחשב גם את התוחלת השנייה, גם כאן יהיו רק 4 איברים בסכום מפני שישנן 4 אפשרויות לשלשות של X,Y,Z עבורן יש הסתברות חיובית:

$$\begin{gathered}\mathrm{E}(XY+XZ+YZ)=(1\cdot 2+1\cdot 3+2\cdot 3)\cdot \mathrm{P}(X=1,Y=2,Z=3)\\ +(2\cdot 1+2\cdot 1+1\cdot 1)\cdot \mathrm{P}(X=2,Y=1,Z=1)\\ +(2\cdot 2+2\cdot 1+2\cdot 1)\cdot \mathrm{P}(X=2,Y=2,Z=1)\\ +(2\cdot 3+2\cdot 2+3\cdot 2)\cdot \mathrm{P}(X=2,Y=3,Z=2)\\ =\tfrac{11}{4}+\tfrac{5}{4}+\tfrac{8}{4}+\tfrac{16}{4}=10\end{gathered}$$

3. מטילים מטבע הוגן 6 פעמים באופן בלתי תלוי. נגדיר X_i־ משתנה מקרי המקבל את הערך 1 אם בהטלה ה־ i התקבל עץ ו־0 אם בהטלה ה־ i התקבל פלי, $i\in\{1,2...6\}$

(א) יהי $S=\sum_{i=1}^{6}X_i$ חשבו את $\mathrm{E}(S)$

(ב) חשבו את $\mathrm{P}(X_1=X_3=X_5=X_6)$

(ג) נגדיר: $Y=X_2\cdot X_4\cdot X_6$ חשבו את פונקצית ההסתברות המשותפת של Y ו־ S. (אפשר לרשום את התשובה כטבלה)

<u>פתרון:</u>

(א) תוחלת של סכום של משתנים מקריים שווה לסכום התוחלות כלומר: $\mathrm{E}(S)=\mathrm{E}\left(\sum_{i=1}^{6}X_i\right)=\sum_{i=1}^{6}\mathrm{E}(X_i)$.
נשים לב ש־ $X_i\sim Ber\left(\frac{1}{2}\right)$ לכל i, ולכן: $\mathrm{E}(X_i)=\frac{1}{2}$ לכל i.
כלומר:

$$\mathrm{E}(S)=\sum_{i=1}^{6}\mathrm{E}(X_i)=\sum_{i=1}^{6}\frac{1}{2}=6\cdot\frac{1}{2}=3$$

(ב) כדי לחשב את ההסתברות המבוקשת, נשתמש בנוסחא $\mathrm{P}(A)=\frac{|A|}{|\Omega|}$.
מותר לנו להשתמש בנוסחא הזו מכיוון שמרחב המדגם שלנו אחיד ־ ההסתברות לכל צירוף של תוצאות אפשריות בארבעת המ`מ היא זהה (מטבעות הוגנים).
Ω היא כל הצירופים האפשריים של 4 תוצאות ולכן $|\Omega|=2^4=16$.
A הוא כל הצירופים שבהם ארבעת התוצאות זהות כלומר $\{1,1,1,1\}$ ו־ $\{0,0,0,0\}$ ולכן $|A|=2$.
לכן: $\mathrm{P}(A)=\frac{2}{16}$

(ג) תחילה, נמצא את התומכים: Y יכול לקבל רק את הערכים $\{0,1\}$, ואילו S יכול לקבל את הערכים $\{0,1,2,3,4,5,6\}$.
נוכל לחשב את ההסתברויות השוליות של Y(שימו לב שהשתמשנו באי התלות בין הטלות המטבע):
בנוסף, נוכל לראות שהערך ש־S מקבל הוא בעצם מספר המטבעות שהראו `עץ`. כלומר, S סופר את מספר הפעמים שיצא `עץ` ב־6 הטלות מטבע ב`ת ושוות הסתברות ולכן: $S\sim Bin\left(6,\frac{1}{2}\right)$.
נשתמש בנוסחת ההסתברות של מ`מ בינומי כדי לחשב את ההסתברויות השוליות של S.
נסדר את החישובים שעשינו בטבלה:

S / Y	0	1	2	3	4	5	6	$\mathrm{P}(Y=y$
0								$\frac{7}{8}=\frac{56}{64}$
1								$\frac{1}{8}=\frac{8}{64}$
$\mathrm{P}(S=s)$	$\binom{6}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^6$	$\binom{6}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^5$	$\binom{6}{2}\left(\frac{1}{2}\right)^6$	$\binom{6}{3}\left(\frac{1}{2}\right)^6$	$\binom{6}{4}\left(\frac{1}{2}\right)^6$	$\binom{6}{5}\left(\frac{1}{2}\right)^6$	$\binom{6}{6}\left(\frac{1}{2}\right)^6$	

נחשב את ההסתברויות המשותפות (כמובן, שנתחיל מהחישובים הפשוטים וננסה להימנע מהחישובים המורכבים בעזרת השלמה של עמודות ושורות בטבלה):
נשים לב ש:

$$\mathrm{P}(S=0,Y=1)=\mathrm{P}(S=1,Y=1)=\mathrm{P}(S=2,Y=1)=0$$

מכיוון ש־Y מקבל את הערך 1 אך ורק אם $X_2=X_4=X_6=1$ ובמצב כזה, לא יכול להיות שהסכום יהיה קטן מ־3.
בנוסף, מתקיים:

$$\mathrm{P}(S=6,Y=0)=0$$

מכיוון שכדי שסכום המ`מ יהיה 6, כולם חייבים לקבל את הערך 1, ובמצב כזה לא יכול להיות שהמכפלה תתאפס.
כדי שיתקיים $\{S=3,Y=1\}$, חייב להתקיים $\{X_1=X_3=X_5=0\,,\,X_2=X_4=X_6=1\}$ ולכן:

$$\mathrm{P}(S=3,Y=1)=\left(\frac{1}{2}\right)^6=\frac{1}{64}$$

כדי שיתקיים $\{S=4,Y=1\}$, חייב להתקיים $\{X_2=X_4=X_6=1\}$ וגם, שבדיוק אחד משלושת ה־ X_i הנותרים יקבל את הערך 1. לכן:

$$\mathrm{P}(S=4,Y=1)=\left(\frac{1}{2}\right)^3\cdot\left(\binom{3}{1}\left(\frac{1}{2}\right)^2\cdot\frac{1}{2}\right)=\frac{3}{64}$$

כדי שיתקיים $\{S=5,Y=1\}$, חייב להתקיים $\{X_2=X_4=X_6=1\}$ וגם, שבדיוק שניים משלושת ה־ X_i הנותרים יקבלו את הערך 1. לכן:

$$\mathrm{P}(S=5,Y=1)=\left(\frac{1}{2}\right)^3\cdot\left(\binom{3}{2}\cdot\frac{1}{2}\cdot\left(\frac{1}{2}\right)^2\right)=\frac{3}{64}$$

נוכל למלא את הטבלה:

$Y \backslash S$	0	1	2	3	4	5	6	$P_Y(y)$
0	$\frac{1}{64}$	$\frac{6}{64}$	$\frac{15}{64}$	$\frac{19}{64}$	$\frac{12}{64}$	$\frac{3}{64}$	0	$\frac{56}{64}$
1	0	0	0	$\frac{1}{64}$	$\frac{3}{64}$	$\frac{3}{64}$	$\frac{1}{64}$	$\frac{8}{64}$
$P_S(s)$	$\frac{1}{64}$	$\frac{6}{64}$	$\frac{15}{64}$	$\frac{20}{64}$	$\frac{15}{64}$	$\frac{6}{64}$	$\frac{1}{64}$	

נוודא שבאמת קיבלנו פונקציית הסתברות כלומר, שכל שורה וכל עמודה מסתכמות להסתברות השולית המתאימה וכן, שכל ההסתברויות השוליות וכל ההסתברויות המשותפות מסתכמות (כל אחת בנפרד) ל־ 1.

4. מטילים קוביה הוגנת n פעמים. נגדיר את שני המשתנים המקריים הבאים:

X־ מספר הפעמים שהתוצאה היא 1

Y־ מספר הפעמים שהתוצאה היא 6

חשבו את השונות המשותפת ומקדם המתאם של X ו־Y

<u>פתרון:</u>

נשים לב ש $X\backsim Bin(n,\frac{1}{6})$ וגם $Y\backsim Bin(n,\frac{1}{6})$. בנוסף, ההתפלגות של $X+Y$ היא גם בינומית: $X+Y\backsim Bin(n,\frac{2}{6})$ (סופרים כמה פעמים יצא 1 או 6 ב־n הטלות בלתי תלויות) מכאן אפשר להסיק:

$$\mathrm{Var}(X)=\mathrm{Var}(Y)=np(1-p)=n\cdot\frac{1}{6}\cdot\frac{5}{6}=\frac{5n}{36}$$

$$\mathrm{Var}(X+Y)=n\cdot\frac{2}{6}\cdot\frac{4}{6}=\frac{8n}{36}$$

הוכחנו טענה על שונות של סכום של משתנים מקריים לפיה:

$$\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)+2\cdot\mathrm{Cov}(X,Y)$$

נציב ונקבל: $\mathrm{Cov}(X,Y)=-\frac{n}{36}$ נחלק בסטיות התקן של X ו־Y ונקבל $\mathrm{Corr}\,(X,Y)=-\frac{1}{5}$

5. נניח כי X_1,X_2 הם משתנים מקריים בלתי תלויים המתפלגים $\mathrm{Bin}(1,1/2)$. נסמן $X=|X_1-X_2|$ ו־$Y=X_1\cdot X_2$.

(א) סמנו את הטענה הנכונה:

i. X,Y הם בלתי תלויים ולכן גם בלתי מתואמים.

ii. X,Y אינם בלתי תלויים, אך הם בלתי מתואמים.

iii. X,Y אינם בלתי מתואמים ולכן גם אינם בלתי תלויים.

iv. X,Y הם שווי התפלגות.

(ב) סמנו את הטענה הנכונה:

i. X,X_1 הם בלתי תלויים ולכן גם בלתי מתואמים

ii. X,X_1 אינם בלתי תלויים, אך הם בלתי מתואמים.

iii. X,X_1 אינם בלתי מתואמים ולכן גם אינם בלתי תלויים

iv. X,X_1 אינם שווי התפלגות.

<u>פתרון:</u>

(א) תשובה ג נכונה.

נמצא את פונקציית ההתפלגות המשותפת של X ו־Y:

נשים לב שהתפלגות $Bin(1,\frac{1}{2})$היא בעצם $Ber(\frac{1}{2})$ ולכן התומכים של X_1 ו־X_2 הם {0,1} :

X הוא ההפרש בערך מוחלט בין X_1 ו־ X_2 ולכן הוא מקבל את הערך 1 כאשר X_1 ו־ X_2 שונים זה מזה ואת הערך 0 כאשר הם זהים.

Y הוא המכפלה של X_1 ו־ X_2 ולכן הוא מקבל את הערך 1 אם שניהם שווים ל־1 ו־0 אחרת.

נחשב את ההסתברויות השוליות:

$$\mathrm{P}(X=1)=\mathrm{P}(\{X_1=1\cap X=0\}\cup\{X_1=0\cap X_2=1\})=\frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{2}$$
$$\mathrm{P}(X=0)=1-\mathrm{P}(X=1)=\frac{1}{2}$$
$$\mathrm{P}(Y=1)=\mathrm{P}(X_1=1\cap X_2=1)=\frac{1}{4}$$
$$\mathrm{P}(Y=0)=1-\mathrm{P}(Y=1)=\frac{3}{4}$$

נשים לב שמתקיים $\mathrm{P}(X=1\cap Y=1)=0$ מכיוון ש־ Yמקבל את הערך 1 רק אם X_1 ו־ X_2 מקבלים את הערך 1 אבל במקרה זה Xיהיה 0.

נרשום את ההסתברויות שחישבנו בטבלה, ונשלים אותה :

Y \ X	0	1	
0	$1/4$	$1/2$	$3/4$
1	$1/4$	0	$1/4$
	$1/2$	$1/2$	

נוכל לראות כי X ו־Yאינם בלתי תלויים, למשל: $\mathrm{P}(X=1,Y=1)=0\neq\mathrm{P}(X=1)\cdot\mathrm{P}(Y=1)=\frac{1}{2}\cdot\frac{1}{4}$

נבדוק האם הם מתואמים על ידי חישוב השונות המשותפת:

$$\mathrm{Cov}(X,Y)=\mathrm{E}(XY)-\mathrm{E}(X)\cdot\mathrm{E}(Y)=-\frac{1}{8}$$

לכן X ו־Y אינם בלתי מתואמים.

(ב) תשובה א נכונה.

נמצא את פונקצית ההתפלגות המשותפת של X ו־ X_1.

פונקציית ההסתברות השולית של X_1 נתונה לנו, ובסעיף הקודם חישבנו גם את זו של X

נחשב:

$$\mathrm{P}(X_1=1\cap X=1)=\mathrm{P}(X_1=1\cap X_2=0)=\mathrm{P}(X_1=1)\cdot\mathrm{P}(X_2=0)=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$$

המעבר האחרון נובע מאי התלות של X_1 ו־ X_2 כפי שנתון.

נמלא את הטבלה:

Y \ X	0	1	
0	$1/4$	$1/2$	$3/4$
1	$1/4$	0	$1/4$
	$1/2$	$1/2$	

נוכל לראות שמתקיים $\mathrm{P}(X=s\cap X=t)=\mathrm{P}(X=s)\cdot\mathrm{P}(X_1=t)$ לכל s,t
כלומר X ו־ X_1 בלתי תלויים ולכן גם בלתי מתואמים.

6. נתונה פונקצית ההסתברות המשותפת הבאה:

Y \ X	0	1	2	
0	0.03	0.02	0.05	0.1
1	0.09	0.06	0.15	0.3
2	0.18	0.12	0.3	0.6
	0.3	0.2	0.5	

האם X ו־ Y ב`ת? אם כן, הראו שהשוויון: $\mathrm{E}(XY)=\mathrm{E}(X)\mathrm{E}(Y)$ מתקיים

<u>פתרון:</u>

נוכל לראות שמתקיים $\mathrm{P}(X=s\cap Y=t)=\mathrm{P}(X=s)\cdot\mathrm{P}(Y=t)$ לכל s,t כלומר X ו־ Y בלתי תלויים

נראה שהשוויון: $\mathrm{E}(XY)=\mathrm{E}(X)\mathrm{E}(Y)$ מתקיים:

$$\begin{gathered}\mathrm{E}(X)=0\cdot0.3+1\cdot0.2+2\cdot0.5=1.2\\ \mathrm{E}(Y)=0\cdot0.1+1\cdot0.3+2\cdot0.6=1.5\\ \mathrm{E}(XY)=0\cdot0\cdot0.03+0\cdot1\cdot0.02+0\cdot2\cdot0.05...2\cdot2\cdot0.3=1.8\\ \Downarrow\\ 1.2\cdot1.5=1.8\end{gathered}$$

7. משקיע מעוניין לשהקיע סכום כסף מסוים בשתי מניות X_1 ו־ X_2 בעלות תוחלת רווח $\mu_1 = 0.075$ ו־ $\mu_2 = 0.1$, שוניות $\sigma_1^2 = 0.01$ ו־ $\sigma_2^2 = 0.008$ ושונות משותפת: $\mathrm{Cov}(X_1, X_2) = -0.0009$. המשקיע שונא סיכון ולכן מעוניין למזער את שונות תיק ההשקעות שלו. מצא את תיק ההשקעות מהצורה:

$$X = aX_1 + (1-a)X_2$$

אשר ממזער את שונות התשואה על ההשקעה, כלומר :

$$a^* = \underset{a}{\mathrm{argmin}}\mathrm{Var}(aX_1 + (1-a)X_2)$$

מהן תוחלת ושונות התשואה של התיק שנבחר?

פתרון:

על פי הטענה שראיתם בפרק זה

$$a^* = \frac{\mathrm{Var}(X_2) - \mathrm{Cov}(X_1, X_2)}{\mathrm{Var}(X_1) + \mathrm{Var}(X_2) - 2\mathrm{Cov}(X_1, X_2)}$$

במקרה זה נקבל:

$$a^* = \frac{0.008 + 0.0009}{0.01 + 0.008 + 2 \cdot 0.0009} = 0.202$$

תוחלת התשואה של התיק הנבחר תהיה:

$$\mathrm{E}(X) = a^* \cdot \mu_1 + (1-a^*)\mu_2 = 0.202 \cdot 0.1 + (1-0.202) \cdot 0.075 = 0.08$$

והשונות:

$$\mathrm{Var}(X) = {a^*}^2 \cdot \mathrm{Var}(X_1) + (1-a^*)^2 \cdot \mathrm{Var}(X_2) + 2 \cdot a^* \cdot (1-a^*) \cdot \mathrm{Cov}(X_1, X_2)$$
$$= 0.202^2 \cdot 0.01 + (1-0.202)^2 \cdot 0.008 + 2 \cdot 0.202 \cdot (1-0.202) \cdot (-0.0009) = 0.00052$$

37 אי־שוויונים

1. רוצים לבדוק את אמינות הרישום של שני מיליון תינוקות שנולדו בניו יורק. מתוך שני מליון נדגמו $1,026,000$ בנים. בהנחה כי ההסתברות לבן ולבת שווה וכי הלידות הן בלתי תלויות. האם התוצאה הזו סבירה? כלומר, מהי ההסתברות שמספר הבנים שנדגמו הוא לפחות $1,026,000$? תנו חסם להסתברות זאת בעזרת אי שיוויון מרקוב וצ'יבצ'ב והשוו ביניהם.

פתרון:

נגדיר X משתנה מקרי הסופר את מספר הבנים מתוך שני מיליון התינוקות. אם ההסתברות לבן היא חצי, והלידות הן בלתי תלויות אז $X \backsim Bin(2,000,000, 0.5)$ ולכן: $\mathrm{E}(X) = 1,000,000$ ו־$\mathrm{Var}(X) = 500,000$

על פי אי שוויון מרקוב:

$$\mathrm{P}(X > 1026000) \leq \frac{1,000,000}{1026000} \cong 0.974$$

על פי אי שוויון צ'ביצ'ב:

$$\mathrm{P}(X \geq 1,026,000) = \mathrm{P}(X - 1,000,000 \geq 1,026,000 - 1,000,000) = \mathrm{P}(X - 1,000,000 \geq 26,000)$$
$$= \tfrac{1}{2}\mathrm{P}(\mid X - \mathrm{E}(X) \mid \geq 26,000) \leq \tfrac{1}{2}\tfrac{500,000}{26,000^2} = \tfrac{1}{2704} \cong 0.0004$$

אי שוויון צ'ביצ'ב במקרה זה מעניק לנו חסם אינפורמטיבי מזה של מרקוב, קיבלנו שההסתברות לתוצאה זו היא לכל היותר 0.0004 ולכן סביר לחשוד למשל, שיש טעות במדגם או שההנחה שההסתברות שווה לבנים ולבנות אינה נכונה.

2. לחברת ביטוח יש 10,000 מבוטחים שרכשו ביטוח רכב, כל אחד מהם משלם פרמיה שנתית בגובה $r > 0$. סכום התביעה השנתי של מבוטח מהחברה הוא משתנה מקרי בעל תוחלת של 240 ש`ח וסטיית תקן של 800 ש`ח. נניח כי התביעות של המבוטחים ב`ת. בעזרת החוק החלש של המספרים הגדולים רשמו קירוב לרווח הממוצע של החברה עבור כל מבוטח.

פתרון:

נסמן ב־X_iאת התביעה השנתית של מבוטח $i \in \{1, 2...10000\}$וב־$Y_i = r - X_i$את הרווח של החברה מהמבוטח ה־i. נתון כי $\mathrm{E}(X_i) = 240$ ולפי החוק החלש של המספרים הגדולים, התביעה הממוצעת של מבוטח שואפת לתוחלת ומכיוון שיש מספר גדול של מבוטחים (10000) , ניתן לקבל קירוב לרווח הממוצע של החברה ממבוטח:

$$\overline{Y_{10000}} = \frac{1}{10000}\sum_{i=1}^{10000} Y_i = \frac{1}{10000}\sum_{i=1}^{10000}(r - X_i) = r - \overline{X_{10000}} \cong r - 240$$